DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

SIMULTANEOUS EQUATIONS MODELS FOR DUMMY ENDOGENOUS VARIABLES: A GAME THEORETIC FORMULATION WITH AN APPLICATION TO LABOR FORCE PARTICIPATION

Paul A. Bjorn and Quang H. Vuong

SOCIAL SCIENCE WORKING PAPER 537

July 1984

ABSTRACT

A game theoretic approach for formulating simultaneous equations models for dummy endogenous variables is proposed and applied to a study of husband/wife labor force participation. A distinctive feature of our approach is that the simultaneous model is derived from optimizing behavior as an outcome of a game between two players. The equilibrium concept used is that of Nash. In addition, we show that the logical consistency conditions implied by usual simultaneous equation models with structural shift actually rules out simultaneity for the problem we consider; in our model, no logical consistency conditions are implied on the parameters.

SIMULTANEOUS EQUATIONS MODELS FOR DUMMY ENDOGENOUS VARIABLES: A GAME THEORETIC FORMULATION WITH AN APPLICATION TO LABOR FORCE PARTICIPATION*

Paul A. Bjorn and Quang H. Vuong
California Institute of Technology

1. INTRODUCTION

Over the last decades, economists have been increasingly interested in studying economic decisions involving choice among a finite number of alternatives. For instance, frequently analyzed behavioral phenomena are decisions on labor force participation, travel mode, and brands of commodity purchases. A reason for such a trend may lie in the increasing availability and quality of large microdata sets. Since the behavioral phenomena of interest were qualitative in nature, new statistical models such as the by-now well-known logit model were introduced in econometrics (McFadden (1974), Nerlove and Press (1973, 1976)).

Following the development of the standard linear simultaneous equations model (Koopmans and Hood (1953)), the literature on discrete variables models has rapidly evolved in simultaneous modeling. In particular, the first simultaneous models that were proposed were directly issued from the standard linear simultaneous equations model. Specifically, the models were formulated in terms of a linear simultaneous equations model in latent continuous variables from which the observed qualitative variables were generated using a dichotomization such as $Y_{1t} = 1$ if $Y^*_{1t} > 0$, and $Y_{1t} = 0$ if $Y^*_{1t} \leq 0$ (see e.g., Maddala and Lee (1976), Nelson and Olson (1978)).

More recently, starting with Amemiya (1974) and Heckman (1978), a new generation of simultaneous models for qualitative/truncated variables was introduced where the underlying simultaneous equations models were formulated in terms of both latent continuous variables and observed qualitative/truncated variables. These models are often called simultaneous models with structural shift. As is well-known, however, a major difficulty that arises with these latter models is that they require the parameters to satisfy some logical consistency conditions in order that the models be statistically meaningful (see, e.g., Heckman (1978), Gourieroux, Laffont, and Monfort (1980), and Schmidt (1981)).

With some rare exceptions (Waldman (1981)), the logical consistency conditions implied by the simultaneous models with structural shift do not have an economic interpretation. This fact explains the few applications of these models in economics. Moreover, as Schmidt (1981) has pointed out, when the exogenous variables are qualitative, the associated logical consistency conditions imply some recursivity in the simultaneous equation models, a feature that casts some doubts on the true simultaneous nature of these models. Though Maddala (1983, Sections 5.7 and 5.8) has warned against the mechanistic formulation of simultaneous models with latent continuous variables, and has argued that in many cases an alternative model without logical consistency conditions can be specified, the previous models are nevertheless the only ones available in the literature that have both latent continuous variables and observed qualitative

variables in the equations.

The purpose of this paper is to present an alternative approach for formulating simultaneous equations models for qualitative endogenous variables. We believe our approach is a first step in integrating game theory into discrete choice econometric modeling. For the simple model that we shall propose, both latent continuous variables and observed dichotomous variables will appear in the equations. Contrary to earlier simultaneous models, however, our approach will not impose logical consistency constraints on the parameters.

Second, a distinctive feature of our approach is that the simultaneous model will be derived from optimizing behavior. To be more explicit, our model will be obtained as an outcome of a game between two individuals/players, and the statistical model will be generated by invoking the random utility framework introduced by McFadden (1974, 1981).

Third, as an empirical application of our approach, we will present a study of the joint decision of a husband and wife whether or not to participate in the labor force. We feel that our analysis constitutes a contribution to this problem because we explicitly model the behavior of a married couple in a game theoretic framework. Previous work on this subject has either taken the husband's decision whether or not to work as exogenous (see, e.g., Heckman (1974), Heckman and Macurdy (1980)), or specified the labor supply of a husband and wife from the outcome of a joint utility function, implicitly ignoring, the problem of aggregation of preferences (see, e.g., Ashenfelter and Heckman (1974), Cotterman (1981), and Gronau (1973)).

The paper is organized as follows. In Section 2, we shall show that the logical consistency conditions implied by simultaneous equation models with structural shift actually rules out simultaneity for the simple problem that we consider. Since simultaneity is an inherent feature of the problem, it follows that these models are inappropriate. Then in Section 3, our game-theoretic approach will suggest a natural rule for generating the observed dichotomous variables, other than the dichotomization rule used in the literature up to now. The resulting statistical model is derived, and it is seen that no logical consistency conditions are implied on the parameters. In Section 4, identification and estimation problems are discussed. In Section 5, the empirical example is presented. Section 6 discusses important generalizations, and Section 7 concludes the paper. Proofs of all propositions are found in Appendix A and the construction of data are presented in Appendix B.

2. SIMULTANEOUS EQUATIONS MODELS WITH STRUCTURAL SHIFT: A CRITIQUE

To simplify the discussion, we shall restrict our attention to the case where there are only two qualitative endogenous variables. To simplify further, we shall assume that these variables are dichotomous. As an illustration, it will be convenient to consider the decision to participate in the labor force by a husband and his wife:

$$
\begin{aligned}
Y_h &= 1 \quad \text{if the husband works,} \\
&= 0 \quad \text{otherwise,} \\
Y_w &= 1 \quad \text{if the wife works,} \\
&= 0 \quad \text{otherwise.}
\end{aligned}
$$

(The subscript indexing the observations is omitted in this section and the following one.)

Most empirical studies on husband-wife labor force participation have postulated the maximization of a joint objective function. Following the classical tradition in economics, we shall nevertheless postulate that each individual maximizes his/her own utility function.

The husband and wife decisions need not, however, be independent since the utility derived by each individual naturally depends on the action taken by the other. Let $\tilde{U}_h(i,j)$ be the utility that the husband derives from taking action i if his wife takes action j where i = 1 if the husband works and 0 otherwise, and j=1 if the wife works and 0 otherwise. The utility $\tilde{U}_w(j,i)$ derived by the wife from taking action j if her husband takes action i is defined similarly.

To generate the observed dichotomous variables Y_h and Y_w, the rule that has been used in the literature on simultaneous equations models for dummy endogenous variables is summarized in the following assumption.

ASSUMPTION A1:

$$Y_h = 1 \quad \text{if } Y_h^* > 0,$$
$$\quad = 0 \quad \text{otherwise},$$
$$Y_w = 1 \quad \text{if } Y_w^* > 0,$$
$$\quad = 0 \quad \text{otherwise}$$

where:

$$Y_h^* = Y_w[\tilde{U}_h(1,1) - \tilde{U}_h(0,1)] + (1 - Y_w)[\tilde{U}_h(1,0) - \tilde{U}_h(0,0)],$$
$$Y_w^* = Y_h[\tilde{U}_w(1,1) - \tilde{U}_w(0,1)] + (1 - Y_h)[\tilde{U}_w(1,0) - \tilde{U}_w(0,0)].$$

The rationale for this model is the following. If the wife works ($Y_w = 1$), then the husband works ($Y_h = 1$) if and only if $\tilde{U}_h(1,1) \geq \tilde{U}_h(0,1)$.[1] On the other hand, if the wife does not work ($Y_w = 0$), the husband works if and only if $\tilde{U}_h(1,0) \geq \tilde{U}_h(0,0)$. Combining these two conditions, it follows that $Y_h = 1$ if and only if $Y_h^* > 0$ where Y_h^* is defined as above. The decision rule for Y_w is similarly derived.

Following McFadden (1974, 1981), the utilities $\tilde{U}_h(i,j)$ and $\tilde{U}_w(j,i)$ are treated as random, and decomposed into deterministic components and random components:

$$\tilde{U}_h(i,j) = U_h(i,j) + \eta_h(i,j),$$
$$\tilde{U}_w(j,i) = U_w(j,i) + \eta_w(j,i).$$

To complete the specification of the statistical model, assumptions must be made on the unobserved random components $\eta_h(i,j)$ and $\eta_w(j,i)$. From the expressions for Y_h^* and Y_w^*, it is clear that only the differences $\eta_h(1,1) - \eta_h(0,1)$, $\eta_h(1,0) - \eta_h(0,0)$, $\eta_w(1,1) - \eta_w(0,1)$, and $\eta_w(1,0) - \eta_w(0,0)$ are relevant. We make the following simplifying assumption:

ASSUMPTION A2:

$$\eta_h(1,1) - \eta_h(0,1) = \eta_h(1,0) - \eta_h(0,0) = \varepsilon_h,$$
$$\eta_w(1,1) - \eta_w(0,1) = \eta_w(1,0) - \eta_w(0,0) = \varepsilon_w,$$

where $(\varepsilon_h, \varepsilon_w)$ are normally distributed with zero means, unit variances, and correlation ρ.

To simplify further, we shall assume that the difference in utility that the husband derives from working versus not working, when the wife works, differs only by a constant β_h from the utility he derives from working versus not working when the wife does not work. A similar simplifying assumption is made for the wife. In other words, we rule out second-order effects. Formally, we have:

ASSUMPTION A3:

$$U_h(1,1) - U_h(0,1) = \beta_h + U_h(1,0) - U_h(0,0) \equiv \beta_h + \Delta_h,$$
$$U_w(1,1) - U_w(0,1) = \beta_w + U_w(1,0) - U_w(0,0) \equiv \beta_w + \Delta_w.$$

Using Assumptions A2 and A3 in the expressions for Y_h^* and Y_w^*, it follows that:

$$Y_h^* = \Delta_h + \beta_h Y_w + \varepsilon_h, \tag{1}$$
$$Y_w^* = \Delta_w + \beta_w Y_h + \varepsilon_w. \tag{2}$$

Given the previous assumptions, our model is a simultaneous equations model with structural shift (Heckman (1978), Schmidt (1981)). From Schmidt's condition 12.6, it follows that, for the model to be statistically meaningful, the following logical consistency condition must hold:

$$\beta_h \cdot \beta_w = 0, \tag{3}$$

i.e., either β_h or β_w must be zero.

In the simple problem considered here, the logical consistency condition can readily be interpreted. For instance, "$\beta_h = 0$" means that, apart from the statistical association between ε_h and ε_w, and hence between ε_h and Y_w, the husband's decision on labor force participation does not depend on the actual wife's decision.[2] Thus the logical consistency condition (3) implies that the decision of either one of the individuals must be structurally independent from the decision of the other. In other words, the logical consistency condition associated with the above simultaneous equation model with structural shift introduces structural recursivity in the model. Since simultaneity is inherent in our problem, and since there is no reason to impose a priori that β_h or β_w be zero, it follows that the usual approach for formulating simultaneous equation models with structural shift is inappropriate.

3. AN ALTERNATIVE FORMULATION

As argued by Heckman (1978), an important justification for the use of simultaneous equations models with structural shift is that these models can distinguish structural association from purely statistical association among discrete endogenous variables, while alternative models such as those developed by Goodman (1970), Haberman (1974), and, Nerlove and Press (1976) cannot.[3] The previous section has shown, however, that the corresponding simultaneous model with structural shift is inappropriate in our case because of the implied logical consistency condition.

Though Assumption A1 defines the crucial dichotomization that generates the observed discrete variables Y_1 and Y_2, that assumption has not been questioned in the literature on discrete variables modeling. We shall argue that Assumption A1 is in fact the cause of the problems that are associated with the logical consistency conditions. Our approach relies instead on the following assumption:

ASSUMPTION A1': The observed dichotomous variables (Y_h, Y_w) are Nash Equilibrium (NE) outcomes of a game played between the two individuals.

Since the utility derived by each individual depends on the action of the other, the natural framework is that of a game between the two individuals. In situations other than the one considered here, the utility functions can obviously be replaced by the appropriate payoff functions. Assumption A1' considers the non-cooperative Nash Equilibrium concept, though alternative equilibrium concepts, which may be more appropriate, can be invoked as discussed in Section 6.

Assumption A1' is not, however, sufficient to define how the observed dichotomous variables (Y_h, Y_w) are generated. This is so because in many games, and especially for the particular problem considered here, a Nash Equilibrium may not exist or multiple Nash Equilibria may arise. As seen below, this difficulty will be resolved by increasing the number of parameters.

To determine the NE outcomes of the game, we derive the reaction functions of each individual. Since the outcome space is simply {0,1} X {0,1}, there are only 4 possible reaction functions for each player. These are referred to as H_1, H_2, H_3, H_4 for the husband and W_1, W_2, W_3, W_4 for the wife, as displayed in Figure 1 and Figure 2, respectively. For instance, reaction function H_1 for the husband says that, whether or not the wife works, the husband will always choose not to work.

Given that H_i and W_j are the husband and wife's reaction functions, we can readily find the Nash Equilibrium. Table 1 indicates the Nash Equilibria (or lack thereof) for each of the 16 possible pairs of reaction functions

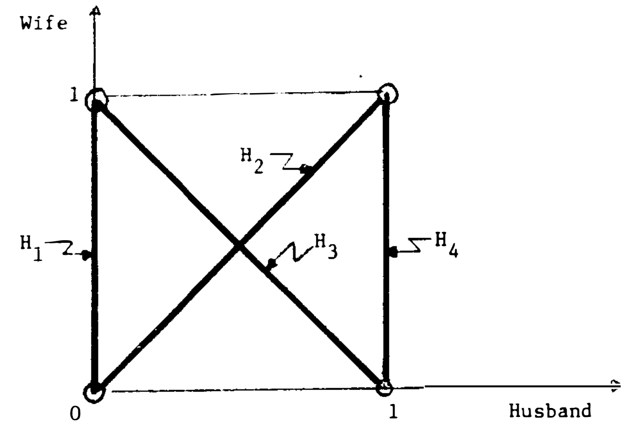

Figure 1: Husband's Reaction Functions

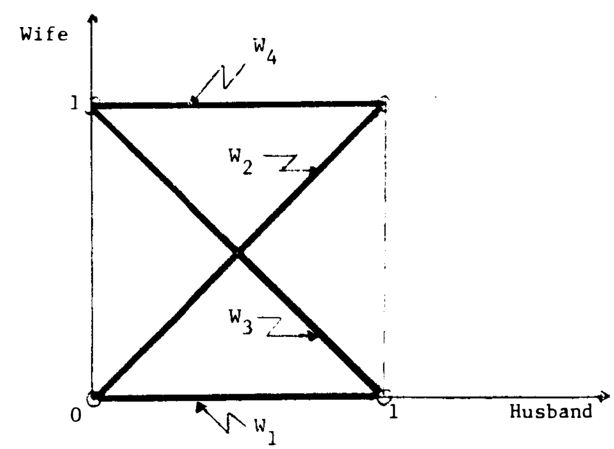

Figure 2: Wife's Reaction Functions

TABLE 1: Nash Equilibria

Husband/Wife	W_1	W_2	W_3	W_4
H_1	(0,0)	(0,0)	(0,1)	(0,1)
H_2	(0,0)	(0,0)&(1,1)	None	(1,1)
H_3	(1,0)	None	(1,0)&(0,1)	(0,1)
H_4	(1,0)	(1,1)	(1,0)	(1,1)

where the first number in the ordered pairs refers to the husband and the second to the wife.

As mentioned earlier, a difficulty arises because of the non-existence or the multiplicity of Nash Equilibria for the pairs (H_2,W_2), (H_2,W_3), (H_3,W_2) and (H_3,W_3). Instead of introducing some arbitrary rules, we shall distribute the probability of occurrence of each of those pairs over the appropriate outcomes according to some weights. The interpretation of the weights will be discussed more fully in Section 6. Let Pr(i,j) be the probability that the random variables Y_h and Y_w take on the values i and j. Using Table 1, it follows that:

$$\begin{aligned} Pr(0,0) &= Pr(H_1 \,\&\, W_1) + Pr(H_1 \,\&\, W_2) + Pr(H_2 \,\&\, W_1) \\ &\quad + a_1Pr(H_2 \,\&\, W_2) + c_1Pr(H_2 \,\&\, W_3) + d_1Pr(H_3 \,\&\, W_2) \end{aligned} \tag{4}$$

$$\begin{aligned} Pr(1,0) &= Pr(H_3 \,\&\, W_1) + Pr(H_4 \,\&\, W_1) + Pr(H_4 \,\&\, W_3) \\ &\quad + b_1Pr(H_3 \,\&\, W_3) + c_2Pr(H_2 \,\&\, W_3) + d_2Pr(H_3 \,\&\, W_2) \end{aligned} \tag{5}$$

$$\begin{aligned} Pr(0,1) &= Pr(H_1 \,\&\, W_3) + Pr(H_1 \,\&\, W_4) + Pr(H_3 \,\&\, W_4) \\ &\quad + b_2Pr(H_3 \,\&\, W_3) + c_3Pr(H_2 \,\&\, W_3) + d_3Pr(H_3 \,\&\, W_2) \end{aligned} \tag{6}$$

$$\begin{aligned} Pr(0,1) &= Pr(H_2 \,\&\, W_4) + Pr(H_4 \,\&\, W_2) + Pr(H_4 \,\&\, W_4) \\ &\quad + a_2Pr(H_2 \,\&\, W_2) + c_4Pr(H_2 \,\&\, W_3) + d_4Pr(H_3 \,\&\, W_2) \end{aligned} \tag{7}$$

where

$$a_1 + a_2 = 1 \quad , \quad b_1 + b_2 = 1 \quad ,$$
$$c_1 + c_2 + c_3 + c_4 = 1 \quad , \quad d_1 + d_2 + d_3 + d_4 = 1 \quad ,$$

all these additional parameters being non-negative.

Each reaction function will, however, occur if some conditions on the random components ε_h and ε_w are satisfied. For instance, given Assumptions A2 and A3, reaction function H_1 arises if and only if:

$$\Delta_h + \varepsilon_h < 0 \quad \text{and} \quad \Delta_h + \beta_h + \varepsilon_h < 0.$$

It is shown in the Appendix that the conditions that must be satisfied by ε_h and ε_w are those given by the following table.

TABLE 2: Conditions for Husband's Reaction Functions

H_1:	$\varepsilon_h < -\Delta_h - \max(0,\beta_h)$
H_2:	$-\Delta_h - \beta_h \leq \varepsilon_h < -\Delta_h$ if $\beta_h \geq 0$; otherwise cannot occur
H_3:	$-\Delta_h \leq \varepsilon_h < -\Delta_h - \beta_h$ if $\beta_h < 0$; otherwise cannot occur
H_4:	$\varepsilon_h \geq -\Delta_h - \min(0,\beta_h)$

Conditions for the wife's reaction functions are essentially the same with the subscript h being replaced by w.

We can now derive the probabilities Pr(i,j) in terms of the unknown parameters. Let $F(a,b,\rho)$ be the c.d.f. evaluated at (a,b) of a bivariate normal distribution with zero means, unit variances, and correlation ρ. Let $I(a,b,c,d,\rho)$ be the integral of the corresponding bivariate density over the range $a \geq \varepsilon_h \geq c$, $b \geq \varepsilon_w \geq d$. Using Equations (4)-(7), we have:

PROPOSITION 1: The probabilities Pr(i,j) are:

$$Pr(0,0) \begin{cases} = F(-\Delta_h, -\Delta_w, \rho) - a_2 I_{++} & \text{if } \beta_h \geq 0,\ \beta_w \geq 0 \\ = F(-\Delta_h, -\Delta_w, \rho) + c_1 I_{+-} & \text{if } \beta_h \geq 0,\ \beta_w < 0 \\ = F(-\Delta_h, -\Delta_w, \rho) + d_1 I_{-+} & \text{if } \beta_h < 0,\ \beta_w \geq 0 \\ = F(-\Delta_h, -\Delta_w, \rho) & \text{if } \beta_h < 0,\ \beta_w < 0 \end{cases} \tag{8}$$

$$Pr(1,0) \begin{cases} = F(\Delta_h, -\Delta_w - \beta_w, -\rho) & \text{if } \beta_h \geq 0,\ \beta_w \geq 0 \\ = F(\Delta_h, -\Delta_w - \beta_w, -\rho) + c_2 I_{+-} & \text{if } \beta_h \geq 0,\ \beta_w < 0 \\ = F(\Delta_h, -\Delta_w - \beta_w, -\rho) + d_2 I_{-+} & \text{if } \beta_h < 0,\ \beta_w \geq 0 \\ = F(\Delta_h, -\Delta_w - \beta_w, -\rho) - b_2 I_{--} & \text{if } \beta_h < 0,\ \beta_w < 0 \end{cases} \tag{9}$$

$$Pr(0,1) \begin{cases} = F(-\Delta_h - \beta_h, \Delta_w, -\rho) & \text{if } \beta_h \geq 0,\ \beta_w \geq 0 \\ = F(-\Delta_h - \beta_h, \Delta_w, -\rho) + c_3 I_{+-} & \text{if } \beta_h \geq 0,\ \beta_w < 0 \\ = F(-\Delta_h - \beta_h, \Delta_w, -\rho) + d_3 I_{-+} & \text{if } \beta_h < 0,\ \beta_w \geq 0 \\ = F(-\Delta_h - \beta_h, \Delta_w, -\rho) - b_1 I_{--} & \text{if } \beta_h < 0,\ \beta_w < 0 \end{cases} \tag{10}$$

$$Pr(1,1) \begin{cases} = F(\Delta_h + \beta_h, \Delta_w + \beta^W, \rho) - a_1 I_{++} & \text{if } \beta_h \geq 0,\ \beta_w \geq 0 \\ = F(\Delta_h + \beta_h, \Delta_w + \beta^W, \rho) + c_4 I_{+-} & \text{if } \beta_h \geq 0,\ \beta_w < 0 \\ = F(\Delta_h + \beta_h, \Delta_w + \beta^W, \rho) + d_4 I_{-+} & \text{if } \beta_h < 0,\ \beta_w \geq 0 \\ = F(\Delta_h + \beta_h, \Delta_w + \beta^W, \rho) & \text{if } \beta_h < 0,\ \beta_w < 0 \end{cases} \tag{11}$$

where

$$\begin{aligned} I_{++} &= I(-\Delta_h, -\Delta_w, -\Delta_h - \beta_h, -\Delta_w - \beta_w, \rho), \\ I_{+-} &= I(-\Delta_h, -\Delta_w - \beta_w, -\Delta_h - \beta_h, -\Delta_w, \rho), \\ I_{-+} &= I(-\Delta_h - \beta_h, -\Delta_w, -\Delta_h, -\Delta_w - \beta_w, \rho), \\ I_{--} &= I(-\Delta_h - \beta_h, -\Delta_w - \beta_w, -\Delta_h, -\Delta_w, \rho). \end{aligned} \tag{12}$$

One can readily check that the four probabilities Pr(i,j) add up to one irrespective of the signs of β_h and β_w, and are continuous with respect to all the parameters. It is then worth noting that the proposed model does not imply any logical consistency constraints on the structural parameters. In addition, Heckman's logical consistency condition (3) can be interpreted in our model.

PROPOSITION 2: Imposing $\beta_h \cdot \beta_w = 0$ is equivalent to imposing that the probability that either one of the four pairs of reaction functions (H_2,W_2), (H_2,W_3), (H_3,W_2), (H_3,W_3) occurs is a priori zero.

In view of Table 1, it follows that the logical consistency condition (3) requires that either the husband's or the wife's action be independent from the action of the other, as discussed in Section 2.

4. IDENTIFICATION AND ESTIMATION

Given the previous expressions for the probabilities Pr(i,j) of the observed dichotomous variables Y_h and Y_w, the log-likelihood function can be written as:

$$\begin{aligned} L &= \sum_t \log Pr_t(Y_{ht}, Y_{wt}) \qquad (13) \\ &= \sum_t [Y_{ht}Y_{wt} \log Pr_t(1,1) + Y_{ht}(1 - Y_{wt}) \log Pr_t(1,0) \\ &+ (1 - Y_{ht})Y_{wt} \log Pr_t(0,1) + (1 - Y_{ht})(1 - Y_{wt}) \log Pr_t(0,0)] \end{aligned}$$

where the subscript t indexes the observations. The probabilities are subscripted by t since Δ_h and Δ_w are in general functions of explanatory variables. We assume:

$$\Delta_{ht} = x_{ht}'\gamma_h \quad \text{and} \quad \Delta_{wt} = x_{wt}'\gamma_w, \qquad (14)$$

where x_{ht} may include characteristics of the t-th household in addition to characteristics of the husband. A similar remark applies to x_{wt}.

The parameters of the model are $(\gamma_h, \beta_h, \gamma_w, \beta_w, \rho)$ together with the weights a's, b's, c's, and d's introduced in Equations (4)-(7). In our first work we constrain a priori these weights to satisfy: $a_1 = a_2 = 1/2$, $b_1 = b_2 = 1/2$, $c_1 = c_2 = c_3 = c_4 = 1/4$, and

$d_1 = d_2 = d_3 = d_4 = 1/4$. In this section and the next, we shall assume that the weights are constrained. Notice that this is equivalent to distributing the probabilities of the four pairs of reaction functions (H_2,W_2), (H_2,W_3), (H_3,W_2), and (H_3,W_3) equally over the appropriate outcomes. In addition to simplifying the maximization of the log-likelihood function, these constraints will simplify identification of the parameters $(\gamma_h, \beta_h, \gamma_w, \beta_w, \rho)$, the problem to which we now turn.

In order to discuss the conditions under which our model is identified, we must first introduce some notation. Define the following partitioned matrix $\tilde{A}$ as

$$\tilde{A} = \left[\bar{X}_h \,\middle|\, D_h^{-1} D_w \bar{X}_w \right]$$ where D_h and D_w are each

block diagonal matrices of order 3T, the t-th block of D_h and D_w given by

$$D_{ht} = \begin{bmatrix} e_t^h & f_t^h & 0 \\ 0 & g_t^h & 0 \\ 0 & 0 & h_t^h \end{bmatrix}, \quad D_{wt} = \begin{bmatrix} e_t^w & 0 & f_t^w \\ 0 & g_t^w & 0 \\ 0 & 0 & h_t^w \end{bmatrix} \quad \text{if} \quad \beta_h > 0,\ \beta_w > 0$$

$$D_{ht} = \begin{bmatrix} e_t^h & 0 & 0 \\ 0 & f_t^h & 0 \\ 0 & -g_t^h & h_t^h \end{bmatrix}, \quad D_{wt} = \begin{bmatrix} e_t^w & 0 & 0 \\ 0 & 0 & f_t^w \\ 0 & -g_t^w & h_t^w \end{bmatrix} \quad \text{if} \quad \beta_h < 0,\ \beta_w < 0$$

$$D_{ht} = \begin{bmatrix} e_t^h & f_t^h & 0 \\ 0 & (g_t^h + f_t^h) & 0 \\ 0 & f_t^h & h_t^h \end{bmatrix}, \quad D_{wt} = \begin{bmatrix} e_t^w & f_t^w & 0 \\ 0 & (g_t^w + f_t^w) & 0 \\ 0 & f_t^w & h_t^w \end{bmatrix} \quad \begin{array}{l} \text{if } \beta_h > 0,\ \beta_w < 0 \\ \text{or } \beta_h < 0,\ \beta_w > 0 \end{array}$$

with

$$\begin{aligned}
e_t^h &= \varphi(\Delta_{ht} + \beta_h)\Phi(\Delta_{wt}^* + \beta_w^* - \rho(\Delta_{ht}^* + \beta_h^*)) \\
f_t^h &= \varphi(\Delta_{ht})\Phi(-\Delta_{wt}^* + \rho\Delta_{ht}^*) \\
g_t^h &= \varphi(\Delta_{ht})\Phi(-\Delta_{wt}^* - \beta_w^* + \rho\Delta_{ht}^*) \\
h_t^h &= -\varphi(\Delta_{ht} + \beta_h)\Phi(\Delta_{wt}^* - \rho(\Delta_{ht}^* + \beta_h^*)) \\
e_t^w &= \varphi(\Delta_{wt} + \beta_w)\Phi(\Delta_{ht}^* + \beta_h^* - \rho(\Delta_{wt}^* + \beta_w^*)) \\
f_t^w &= \varphi(\Delta_{wt})\Phi(-\Delta_{ht}^* + \rho\Delta_{wt}^*) \\
g_t^w &= -\varphi(\Delta_{wt} + \beta_w)\Phi(\Delta_{ht}^* - \rho(\Delta_{wt}^* + \beta_w^*)) \\
h_t^w &= \varphi(\Delta_{wt})\Phi(-\Delta_{ht}^* - \beta_h^* + \rho\Delta_{wt}^*)
\end{aligned}$$

where Φ is the univariate normal c.d.f. and φ is the corresponding p.d.f. A quantity with a "*" means that quantity is divided by the square root of $1 - p^2$. $\bar{X}_h$ and $\bar{X}_w$ are of dimension 3T by $K_h + 1$ and 3T by $K_w + 1$, the t-th blocks given respectively as

$$\begin{bmatrix} 1 & x_{ht}' \\ 0 & x_{ht}' \\ 1 & x_{ht}' \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & x_{wt}' \\ 1 & x_{wt}' \\ 0 & x_{wt}' \end{bmatrix}$$

PROPOSITION 3: The parameters $(\beta_h, \beta_w, \gamma_h, \gamma_w)$ of the model are identified if and only if $\tilde{A}$ is nonsingular.

As can be easily seen, D_h and D_w are nonsingular under all four cases since D_{ht} and D_{wt} are

either triangular matrices or can be made triangular by suitable permutations of rows and columns. By examining matrix $\tilde{A}$ above, it is clear that if $\tilde{A}$ is singular, it will occur extremely rarely for some specific values of the parameters as an artifact of certain explanatory variables (note the exception in Corollary 2).

When compared to the restrictions needed for the identification of linear models, our results are quite surprising. For example, even if x_{ht}' and x_{wt}' are the same, it will occur only rarely that our model is not identified, a sharp contrast to the order conditions which is necessary for the identification of linear models. Moreover, consider the worst possible case for the identification of linear models. Let $x_{ht}' = x_{wt}' \equiv x_t'$, the t-th row of the nonsingular matrix X; further, let $\gamma_h = \gamma_w$ and $\beta_h = \beta_w$. Then it follows that $\Delta_{ht} = \Delta_{wt} \equiv \Delta_t$, $g_t^h = h_t^w \equiv g_t$, and $h_t^h = g_t^w \equiv h_t$. Then we have the following result.

COROLLARY 1: If $x_{ht}' = x_{wt}' \equiv x_t'$, X full column rank, $\gamma_h = \gamma_w$, and $\beta_h = \beta_w$, the model will be identified if and only if the quantity $h_t/(h_t - g_t)$ is a linear combination of x_t, for all t.

Consider the following consequences of the Corollary. First, a linear simultaneous equations model of demand and supply, for example, satisfying the conditions of Corollary 1 will be hopelessly unidentified; the demand and supply curves will be identical. Under the same conditions, our model will be identified almost always. Even more importantly, under the extreme restrictions of Corollary 1, we can meaningfully test for the equality of coefficients between equations since the parameters are identified in this situation.

A further corollary to Proposition 3 provides a necessary condition for identification as shown next. Suppose that constant terms are included in both equations of (14). Define $\Delta_{ht} = \gamma_h^0 + x_{ht}'\tilde{\gamma}_h$ and $\Delta_{wt} = \gamma_w^0 + x_{wt}'\tilde{\gamma}_w$ where γ_h^0 is the coefficient for the husbands' constant term and x_{ht} is a vector of explanatory variables for husband t. γ_w^0 and x_{wt} are defined analogously for the wife. We then have:

COROLLARY 2: If constant terms (γ_h^0 and γ_w^0) are included in both equations, and if $\tilde{\gamma}_h = 0$ and $\tilde{\gamma}_w = 0$, the model is not identified.

Therefore if one intends to estimate the model with both constant terms (γ_h^0 and γ_w^0) included, at least one equation must include further explanatory variables or the model will certainly be not identified.[4]

As a practical implication of the Corollary for estimation, if one includes a constant term in each equation along with one or more additional explanatory variables in at least one of the equations, one or more of the initial values for the parameters associated with these non-constant explanatory variables must be nonzero. Otherwise, the information matrix will be singular at the first iteration, and the optimization cannot be carried out. We now turn to estimation.

The estimation routine we employ is a version of the iterative procedure suggested by Berndt, Hall, Hall, and Hausman (1974). Assuming that the parameters are identified and the a's, b's, c's, and

d's satisfy the a priori weights listed above, maximization is straightforward although the log-likelihood takes four different functional forms according to the signs of β_h and β_w. Since we do not have, for the moment, some initial consistent estimates of the parameters $(\gamma_h, \beta_h, \gamma_w, \beta_w, \rho)$, we provide various initial values for $(\gamma_h, \beta_h, \gamma_w, \beta_w)$ with a grid search over possible values of ρ and iterative until convergence.

5. AN EMPIRICAL EXAMPLE

A. THE MODEL

The following four equations will be used to describe the joint behavior of a representative married couple:

$$W_h^r = Z_h'\gamma_h^r - \beta_h Y_w + \eta_h^r \tag{15}$$

$$W_w^r = Z_w'\gamma_w^r - \beta_w Y_h + \eta_w^r \tag{16}$$

$$W_h^m = X_h'\gamma_h^m + \eta_h^m \tag{17}$$

$$W_w^m = X_w'\gamma_w^m + \eta_w^m \tag{18}$$

Equations (15) and (16) describe the reservation wages, or equivalently, the shadow price of time for the husband and wife, respectively. Note that the wife's decision of whether or not to work, given by the dichotomous variable Y_w, affects the husband's reservation wage in (15). Analogously, the husband's decision of whether or not to work, given by Y_h, affects the wife's reservation wage in (16). Equations (17) and (18) describe the market wages for the husband and the wife, respectively.

Now let the husband's (wife's) reservation wage play the role of the payoff he (she) derives from not working, given by equations (19) and (20) respectively. Similarly, let the husband's (wife's) market wage play the role of the payoff he (she) derives from working, given by equations (21) and (22) respectively. We then have

$$\tilde{U}_h(0,Y_w) = Z_h'\gamma_h^r - \beta_h Y_w + \eta_h^r \tag{19}$$

$$\tilde{U}_w(0,Y_h) = Z_w'\gamma_w^r - \beta_w Y_h + \eta_w^r \tag{20}$$

$$\tilde{U}_h(1,Y_w) = X_h'\gamma_h^m + \eta_h^m \tag{21}$$

$$\tilde{U}_w(1,Y_h) = X_w'\gamma_w^m + \eta_w^m \tag{22}$$

Notice that Assumption A2 on the error terms is naturally satisfied. Moreover, we have:

$$U_h(1,1) - U_h(0,1) = \beta_h + U^h(1,0) - U^h(0,0) = X_h'\gamma_h^m - Z_h'\gamma_h^r + \beta_h \tag{23}$$

$$U_w(1,1) - U_w(0,1) = \beta_w + U^w(1,0) - U^w(0,0) = X_w'\gamma_w^m - Z_w'\gamma_w^r + \beta_w \tag{24}$$

Thus Assumption A3 is also satisfied. In addition, note that in specifying the husband's reservation wage and market wage equations, given by (19) and (21) respectively, it may be the case that certain explanatory variables appear in both equations, implying that the associated coefficient in (23) will be measuring the difference between market and reservation wages. A similar comment holds for the wife.

B. SPECIFICATION OF THE MODEL

We must now specify the set of explanatory variables of the

market wage equations and the reservation wage equations for the husband and wife.[5] Market wages for the husband and wife are specified in (25) and (26) respectively. Reservation wages for the husband and wife are specified in (27) and (28) respectively.

$$\tilde{U}_h(1,\cdot) = W_h^m = \gamma_h^0 + \underset{(+)}{\gamma_h^1}\text{AGEH} + \underset{(+)}{\gamma_h^2}\text{EDUCH} + \underset{(-)}{\gamma_h^3}\text{UNEM} + \underset{(-)}{\gamma_h^4}\text{RACE} + \eta_h^m \tag{25}$$

$$\tilde{U}_w(1,\cdot) = W_w^m = \gamma_w^0 + \underset{(+)}{\gamma_w'}\text{AGEW} + \underset{(+)}{\gamma_w^2}\text{EDUCW} + \underset{(-)}{\gamma_w^3}\text{UNEM} + \underset{(-)}{\gamma_w^4}\text{RACE} + \eta_w^m \tag{26}$$

$$\begin{aligned}\tilde{U}_h(0,Y_w) = W_h^r = \tilde{\gamma}_h^0 &+ \underset{(+)}{\tilde{\gamma}_h^1}\text{AGEH} + \underset{(+)}{\tilde{\gamma}_h^2}\text{EDUCH} + \underset{(-)}{\tilde{\gamma}_h^3}\text{UNEM} + \underset{(+)}{\tilde{\gamma}_h^4}\text{ASSET} + \underset{(?)}{\tilde{\gamma}_h^5}\text{KIDS13} \\ &+ \underset{(?)}{\tilde{\gamma}_h^6}\text{KIDS14} + \underset{(-)}{\tilde{\gamma}_h^7}\text{RACE} - \beta_h Y_w + \eta_h^r\end{aligned} \tag{27}$$

$$\begin{aligned}\tilde{U}_w(0,Y_h) = W_w^r = \tilde{\gamma}_w^0 &+ \underset{(+)}{\tilde{\gamma}_w^1}\text{AGEW} + \underset{(+)}{\tilde{\gamma}_w^2}\text{EDUCW} + \underset{(-)}{\tilde{\gamma}_w^3}\text{UNEM} + \underset{(+)}{\tilde{\gamma}_w^4}\text{ASSET} + \underset{(+)}{\tilde{\gamma}_w^5}\text{KIDS13} \\ &+ \underset{(?)}{\tilde{\gamma}_w^6}\text{KIDS14} + \underset{(-)}{\tilde{\gamma}_w^7}\text{RACE} - \beta_w Y_h + \eta_w^r\end{aligned} \tag{28}$$

where

AGEH	Age of husband
AGEW	Age of wife
EDUCH	Number of years of formal schooling of husband
EDUCW	Number of years of formal schooling of wife
UNEM	Local unemployment rate
RACE	Dummy variable indicating race of married couple; 1 = Black or

Hispanic, 0 otherwise[6]

ASSET Family's annual income other than from wages or salaries [7]

KIDS13 Number of children in family unit 14 years or older

KIDS14 Number of children in family unit 14 years or older

The plus and minus signs under the explanatory variables in equations (25)-(28) indicate the expected impact of each variable in the respective equation. Note that while the wife's decision of whether or not to work has no effect on the husband's market wage, it does affect the husband's reservation wage. Analogously, the husband's decision to work or not affects the wife's reservation wage but not her market wage. From (23) and (24) we now have

$$\begin{aligned} U_h(1,1) - U_h(0,1) &= \beta_h + U_h(1,0) - U_h(0,0) \\ &= (\gamma_h^0 - \tilde{\gamma}_h^0) + (\gamma_h^1 - \tilde{\gamma}_h^1)\text{AGEH} + (\gamma_h^2 - \tilde{\gamma}_h^2)\text{EDUCH} + (\gamma_h^3 - \tilde{\gamma}_h^3)\text{UNEM} \\ &+ (\gamma_h^4 - \tilde{\gamma}_h^4)\text{RACE} - \tilde{\gamma}_h^5\text{ASSETS} - \tilde{\gamma}_h^6\text{KIDS13} - \tilde{\gamma}_H^7\text{KIDS14} + \beta_h \end{aligned} \tag{29}$$

and

$$\begin{aligned} U_w(1,1) - U_w(0,1) &= \beta_w + U_w(1,0) - U_w(0,0) \\ &= (\gamma_w^0 - \tilde{\gamma}_w^0) + (\gamma_w^1 - \tilde{\gamma}_w^1)\text{AGEW} + (\gamma_w^2 - \tilde{\gamma}_w^2)\text{EDUCH} + (\gamma_h^3 - \tilde{\gamma}_h^3)\text{UNEM} \\ &+ (\gamma_w^4 - \tilde{\gamma}_w^4)\text{RACE} - \tilde{\gamma}_w^5\text{ASSETS} - \tilde{\gamma}_w^6\text{KIDS13} - \tilde{\gamma}_w^7\text{KIDS14} + \beta_w \end{aligned} \tag{30}$$

The data used in this study on married couples is from the 1982 wave of the University of Michigan Survey Research Center's Panel Study on Income Dynamics, 1968-1982. The data was restricted to 2020

records for married couples living in the U.S., where both the husband and the wife were able-bodied, neither older than 64 years of age with no nonrelative living in the family unit. A more detailed description of the Panel Study on the selection criteria we used is found in Appendix B.

C. EMPIRICAL RESULTS

From Section 4 it will be recalled that our estimation technique requires that we provide initial values of $(\gamma_h, \beta_h, \gamma_w, \beta_w)$ with a grid search over possible values of ρ and iterate until convergence. Having no strong priors over ρ, we estimated all values from -.9 to +.9 in .1 increments. As can be seen from Table 3, the value of ρ that maximizes the log-likelihood function appears to lie in the interval (-.9, -.7). Therefore we tried all values within this interval at increments of .01. The maximizing value of ρ is -.78, which also appears in Table 3.

Although it may appear surprising that the maximizing value of ρ is not positive, it must be remembered that ρ is not simply a measure of the correlation between omitted variables in the husband's and wife's equations, but is derived from a more complicated relationship given by Assumption A2. A likelihood ratio test of $\rho = 0$ vs. $\rho = -.78$ can be easily constructed from Table 3. Since -2(log-likelihood value for $\rho = 0$ - log-likelihood value for $\rho = -.78$) $\sim \chi_1^2$, we can reject the hypothesis that $\rho = 0$ at the 95 percent level.

Table 5 lists the estimated coefficients and the associated

TABLE 3

Correlation Coefficient (ρ)	Log-likelihood Value
-.9	-1575.01
-.8	-1574.44
-.78	-1574.43
-.7	-1574.64
-.6	-1574.91
-.5	-1575.31
-.4	-1575.71
-.3	-1576.14
-.2	-1576.53
-.1	-1576.92
0.	-1577.30
.1	-1577.69
.2	-1578.06
.3	-1578.44
.4	-1578.78
.5	-1579.09
.6	-1579.48
.7	-1579.77
.8	-1580.23
.9	-1581.13

TABLE 4

$\rho = 0$

Δ_h	Δ_w	Coefficient	Estimate	t-Statistic
		β_h	-1.21	-1.830
		β_w	1.28	2.130
CONSTANT		$(\gamma_h^0 - \tilde{\gamma}_h^0)$	0.886	1.240
AGEH		$(\gamma_h^1 - \tilde{\gamma}_h^1)$	0.018	2.280
EDUCH		$(\gamma_h^2 - \tilde{\gamma}_h^2)$	0.132	3.950
UNEM		$(\gamma_h^3 - \tilde{\gamma}_h^3)$	-0.039	-2.180
RACEH		$(\gamma_h^4 - \tilde{\gamma}_h^4)$	-0.099	-0.720
ASSET		$-\tilde{\gamma}_h^5$	0.383	1.150
KIDS13		$-\tilde{\gamma}_h^6$	-0.161	-2.440
KIDS14		$-\tilde{\gamma}_h^7$	0.111	0.942
	CONSTANT	$(\gamma_w^0 - \tilde{\gamma}_w^0)$	-0.667	-1.150
	AGEW	$(\gamma_w^1 - \tilde{\gamma}_w^1)$	-.009	-2.500
	EDUCW	$(\gamma_w^2 - \tilde{\gamma}_w^2)$	0.032	2.610
	UNEM	$(\gamma_w^3 - \tilde{\gamma}_w^3)$	-0.008	-0.993
	RACEW	$(\gamma_w^4 - \tilde{\gamma}_w^4)$	.398	5.220
	ASSET	$-\tilde{\gamma}_w^5$	-0.015	-2.670
	KIDS13	$-\tilde{\gamma}_w^6$	-0.262	-9.000
	KIDS14	$-\tilde{\gamma}_w^7$	-0.061	-1.290

log-likelihood value = -1577.30

TABLE 5

$\rho = -0.78$

Δ_h	Δ_w	Coefficient	Estimate	t-Statistic
		β_h	-1.010	-1.800
		β_w	2.400	7.780
CONSTANT		$(\gamma_h^0 - \tilde{\gamma}_h^0)$	0.649	0.974
AGEH		$(\gamma_h^1 - \tilde{\gamma}_h^1)$	0.018	2.170
EDUCH		$(\gamma_h^2 - \tilde{\gamma}_h^2)$	0.142	4.420
UNEM		$(\gamma_h^3 - \tilde{\gamma}_h^3)$	-0.038	-2.090
RACEH		$(\gamma_h^4 - \tilde{\gamma}_h^4)$	-0.157	-1.100
ASSET		$-\tilde{\gamma}_h^5$	0.383	1.240
KIDS13		$-\tilde{\gamma}_h^6$	-0.085	-1.500
KIDS14		$-\tilde{\gamma}_h^7$	0.172	1.440
	CONSTANT	$(\gamma_w^0 - \tilde{\gamma}_w^0)$	-1.760	-5.050
	AGEW	$(\gamma_w^1 - \tilde{\gamma}_w^1)$	-0.010	-2.790
	EDUCW	$(\gamma_w^2 - \tilde{\gamma}_w^2)$	0.030	2.530
	UNEM	$(\gamma_w^3 - \tilde{\gamma}_w^3)$	-0.005	-0.622
	RACEW	$(\gamma_w^4 - \tilde{\gamma}_w^4)$	0.412	5.550
	ASSET	$-\tilde{\gamma}_w^5$	-0.014	-2.640
	KIDS13	$-\tilde{\gamma}_w^6$	-0.250	-8.610
	KIDS14	$-\tilde{\gamma}_w^7$	-0.062	-1.320

log-likelihood value = -1574.43

t-statistics for the case $\rho = -.78$. A comparison of Table 4, the estimated coefficients and t-statistics for the case of $\rho = 0$, with Table 5 indicates that the signs of the estimated coefficients are quite robust with respect to ρ. From Table 5 we see that both β_h and β_w are significantly different from zero (β_h only at the 90 percent level), providing evidence that the husband's decision whether or not to work depends on the wife's decision and vice versa. From (15) we see that the negative estimated coefficient of β_h from Table 5 implies that if the wife works, the husband's reservation wage increases as expected. A priori, we would also expect the estimate of β_w to be negative also; we offer no explanation for the disconcerting result that β_w is positive.

Looking again at Table 5, we see that most of the coefficients explaining the wife's decision whether or not to work are in agreement with our expectations and are highly significant. For example, family income from other than wages and salaries (ASSET) and the number of children under the age of 13 (KIDS13) have the expected effect of increasing the wage at which a wife is willing to accept work outside the home.[8] Consistent also with our expectations is the indeterminant effect of number of children 14 years or older (KIDS14) on the wife's reservation wage. While the mother may hesitate seeking outside employment if she has young teenagers, she may feel compelled to supplement the family income for current and future college expenses. The estimated positive coefficient on the female race dummy (RACE) indicates that women of racial minorities, on average, can command a higher market wage than they are willing to accept to enter the labor market; that is, minority women are on average worth more in the marketplace than they think they are worth. While one may interpret this result as saying that racial discrimination in the marketplace is not as widespread as minority women are led to believe, the discrepancy between reservation wage and market wage may be a reflection of past discrimination. The coefficients on the wives' age (AGEW) and education (EDUCW) are also consistent with our prior expectations; although an increase in either age or education should increase the wife's market wage, both should also increase the reservation wage.

Turning next to the variables used to explain the husband's decision of whether or not to work, we see that while a number of the coefficients are insignificant, the variables to which we attached strong priors appear to be significant. For example, the coefficients attached to both the husband's age (AGEH) and the local unemployment rate (UNEM) measure the difference between the husband's market wage and his reservation wage; if the husband is behaving logically, both of these estimated coefficients should be close to zero, which they are. The positive estimated coefficient on the husband's level of education (EDUCH) is surprising; one would a priori expect that an increase in education would raise equally the husband's market wage and reservation wage. One possible explanation for this surprising result is that additional education is viewed by many as a consumption good rather than an investment good.[9]

6. GENERALIZATIONS

First, an obvious generalization of our model is to allow the weights (a_h, b_i, c_j, d_k) to depend on some explanatory variables. For instance:

$$\log a_{1t} = \mu_t^a + z_t'\delta^a$$

where μ_t^a is a normalizing parameter so that $a_{1t} + a_{2t} = 1$ and z_t are some household characteristics. (The logit functional form is chosen so that $0 < a_{1t} < 1$ and $0 < a_{2t} < 1$.) The simultaneous model hence obtained contains our previous model which assumes that the weight parameters are constant across observations, or equivalently that the parameters $\delta^a, \delta^b, \delta^c$, and δ^d are zero, with the exception of those parameters that are associated with the constant term. It follows that this assumption can be tested. Another interesting test would be to see if $\delta^a = \delta^b = \delta^c = \delta^d = 0$ (including the constant term). Indeed, if this hypothesis is not rejected, then the data would support the idea that, when there are two or no Nash Equilibria, these two Nash Equilibria or all the four possible outcomes are respectively equally likely as we assumed.

Our second generalization, which is related to the first one, deals with the interpretation to be given to the weights. In the previous developments, we have restricted ourselves to Nash Equilibria in pure strategies. A broader class of strategies is in general considered in game theory, namely the class of mixed strategies (see, e.g., Owen (1982)). In this more general approach, randomness enters into the statistical model for two reasons which are (i) the random nature of the sampling (or equivalently the econometrician ignorance), and (ii) the randomization of the strategies. How can the simultaneous model proposed above be interpreted when mixed strategies are allowed?

First, it is worth noting that in 12 out of 16 cases, a unique Nash Equilibrium exists (see Table 1). Moreover, it can be shown that when there is a unique Nash Equilibrium in pure strategies, then there are no Nash Equilibrium in mixed strategies (see Appendix, Lemma 1). It follows that for those 12 cases, restricting ourselves to pure strategies is irrelevant.

On the other hand, when there are two Nash Equilibria, which is our case when the husband and wife reaction functions are (H_2,W_2) or (H_3,W_3), then a mixed strategy Nash Equilibrium may exist. However, it can be shown that if this is so, then the (expected) utilities derived by each individual when playing the NE mixed strategies are strictly less than the utilities they would derive by playing at least one of the two available NE pure strategies (see Appendix, Lemma 2). This provides a justification for distributing, as we did, the probability of (H_2,W_2), say, over the outcomes (1,1) and (0,0) only (see Table 1, and Equations (4) and (7)).

The third possibility is when there are no Nash Equilibria in pure strategies. This occurs when the husband and wife's reactions functions are (H_2,W_3) or (H_3,W_2). In this case, it can be shown that there exists a unique Nash Equilibrium in mixed strategies (see

Appendix, Lemma 3). Let (q_{h0},q_{h1}), where $q_{h0} + q_{h1} = 1$, characterizes the randomization between not working and working for the husband; (q_{w0},q_{w1}) is similarly defined. Then, the weights c_1, c_2, c_3, and c_4 defined in Section 3 can be interpreted as:

$$c_1 = q_{h0}q_{w0} \quad , \quad c_2 = q_{h1}q_{w0} \quad ,$$
$$c_3 = q_{h0}q_{w1} \quad , \quad c_4 = q_{h1}q_{w1} \quad .$$

A similar interpretation holds for the weights d_1, d_2, d_3, d_4 associated with the pair (H_3,W_2). It follows that the model proposed in Section 3 can be thought of as being a model in which the weights c's and d's are left unconstrained. Of course, a tighter specification of the model can be derived. Indeed, since the Nash Equilibrium in mixed strategies is unique, it follows that the q's and hence the c's (or d's) are functions of the structural parameters $(\gamma^h,\beta^h,\gamma^w,\beta^w,\rho)$.

7. CONCLUSION

In this paper, we presented an alternative approach for formulating simultaneous equations models for qualitative endogenous variables. Contrary to earlier simultaneous models, our model does not require any logical consistency conditions on the parameters. In addition, a distinctive feature of our approach is that the simultaneous model is derived from optimizing behavior within the random utility framework.

Our approach also emphasizes the role of the equilibrium concept used in order to define the process generating the observed variables Y_h and Y_w. The previous section has shown that the proposed model can readily allow Nash Equilibria in mixed strategies. Other equilibrium concepts are, however, available. For instance, one may consider that one player is moving first in which case a Stackelberg Equilibrium concept may be used (see Vuong (1982)). Alternatively, one may assume that the individuals play cooperatively so as to maximize a joint utility or payoff function. Though these latter models essentially remove simultaneity, it would be interesting to devise statistical procedures that discriminate between these various models. Though obviously important work remains, we believe that our proposed model is a first step towards integrating results in game theory and discrete choice modeling.

Finally, we presented an empirical application of the joint decision of a husband and wife whether or not to participate in the labor force. We feel that our study constitutes a contribution to this problem because we explicitly model the behavior of a married couple in a game theoretic framework, whereas previous work on this subject has either taken the husband's decision as exogenous or specified the labor supply decisions of a husband and wife from the outcome of a joint utility function. Although our model is quite complicated, as can be seen from the programming requirements indicated in Appendix A, we were heartened to get such strong results, both in terms of correct signs on coefficients in which we had strong priors and highly significant t-statistics.

APPENDIX A

a. Conditions for Husband's Reaction Functions

Using Figure 1, reaction function H_1 can be characterized by the following two conditions: $\tilde{U}_h(1,0) - \tilde{U}_h(0,0) < 0$ and $\tilde{U}_h(1,1) - \tilde{U}_h(0,1) < 0$. Using Assumptions 1 and 2 these conditions are equivalent to $\varepsilon_h < - \Delta_h$ and $\varepsilon_h < - \Delta_h - \beta_h$, respectively, which can be combined to give $\varepsilon_h < - \Delta_h - \max(0,\beta_h)$.

Reaction function H_2 is characterized by $\tilde{U}_h(1,0) - \tilde{U}_h(0,0) < 0$ and $\tilde{U}_h(1,1) - \tilde{U}_h(0,1) \geq 0$, which are equivalent to $\varepsilon_h < - \Delta_h$ and $\varepsilon_h \geq - \Delta_h - \beta_h$, respectively. When combined they give the result in the text.

Reaction function H_3 is characterized by $\tilde{U}_h(1,1) - \tilde{U}_h(0,1) < 0$ and $\tilde{U}_h(1,0) - \tilde{U}_h(0,0) \geq 0$. Using Assumptions 1 and 2 these conditions are equivalent to $\varepsilon_h < - \Delta_h - \beta_h$ and $\varepsilon_h \geq - \Delta_h$, respectively. When combined, we get the result in the text.

Reaction function H_4 is characterized by $\tilde{U}_h(1,0) - \tilde{U}_h(0,0) \geq 0$ and $\tilde{U}_h(1,1) - \tilde{U}_h(0,1) \geq 0$, which are equivalent to $\varepsilon_h \geq - \Delta_h$ and $\varepsilon_h \geq - \Delta_h - \beta_h$, respectively, which when combined give $\varepsilon_h \geq - \Delta_h - \min(0,\beta_h)$.

The wife's reaction functions are derived identically.

b. PROOF OF PROPOSITION 1: From Table 2 and the conditions for the wife's reaction functions, it follows that some reaction functions cannot occur according to the signs of β_h and β_w. For instance, when $\beta_h \geq 0$ and $\beta_w \geq 0$, the reaction functions H_3 and W_3 cannot occur. From Equations (4)-(7) it follows that, when $\beta_h \geq 0$ and $\beta_w \geq 0$, we have:

$$Pr(0,0) = Pr(H_1,W_1) + Pr(H_1,W_2) + Pr(H_2,W_1) + a_1 Pr(H_2,W_2),$$
$$Pr(1,0) = Pr(H_4,W_1),$$
$$Pr(0,1) = Pr(H_1,W_4),$$
$$Pr(1,1) = Pr(H_2,W_4) + Pr(H_4,W_2) + Pr(H_4,W_4) + a_2 Pr(H_2,W_2).$$

Similarly, when $\beta_h \geq 0$ and $\beta_w < 0$, the reaction functions H_3 and W_2 never occur so we have:

$$Pr(0,0) = Pr(H_1,W_1) + Pr(H_2,W_1) + c_1 Pr(H_2,W_3),$$
$$Pr(1,0) = Pr(H_4,W_1) + Pr(H_4,W_3) + c_2 Pr(H_2,W_3),$$
$$Pr(0,1) = Pr(H_1,W_3) + Pr(H_1,W_4) + c_3 Pr(H_2,W_3),$$
$$Pr(1,1) = Pr(H_2,W_4) + Pr(H_4,W_2) + Pr(H_4,W_4) + c_4 Pr(H_2,W_3).$$

When $\beta_h < 0$ and $\beta_w \geq 0$, the reaction functions H_2 and W_3 never occur. Thus:

$$Pr(0,0) = Pr(H_1,W_1) + Pr(H_1,W_2) + d_1 Pr(H_3,W_2),$$
$$Pr(1,0) = Pr(H_3,W_1) + Pr(H_4,W_1) + d_2 Pr(H_3,W_2),$$
$$Pr(0,1) = Pr(H_1,W_4) + Pr(H_3,W_4) + d_3 Pr(H_3,W_2),$$
$$Pr(1,1) = Pr(H_4,W_2) + Pr(H_4,W_4) + d_4 Pr(H_3,W_2).$$

Finally, when $\beta_h < 0$ and $\beta_w < 0$, the reaction functions H_2 and

Areas Defining Pr(0,0), Pr(1,0), Pr(0,1), Pr(1,1)

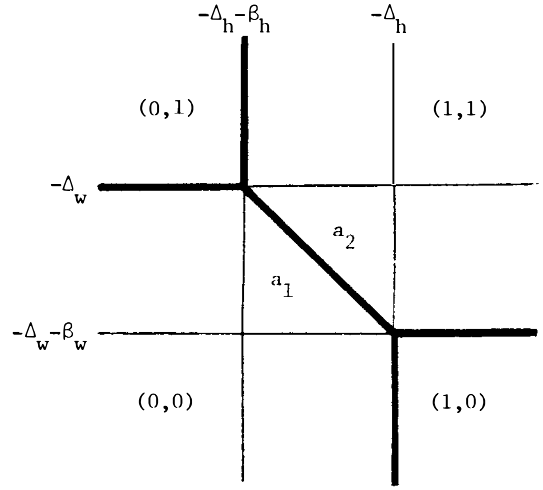

Fig. 3a : $\beta_h \geq 0$, $\beta_w \geq 0$

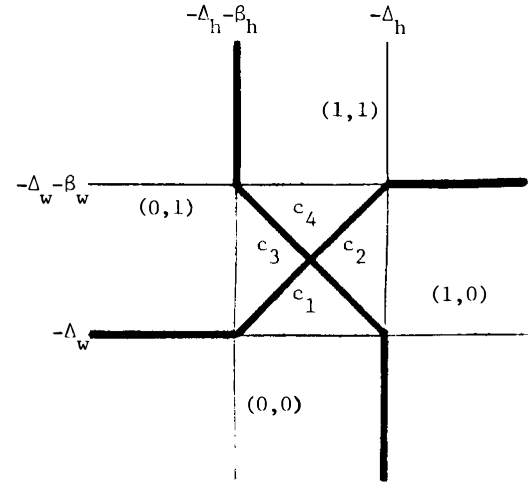

Fig. 3b : $\beta_h \geq 0$, $\beta_w < 0$

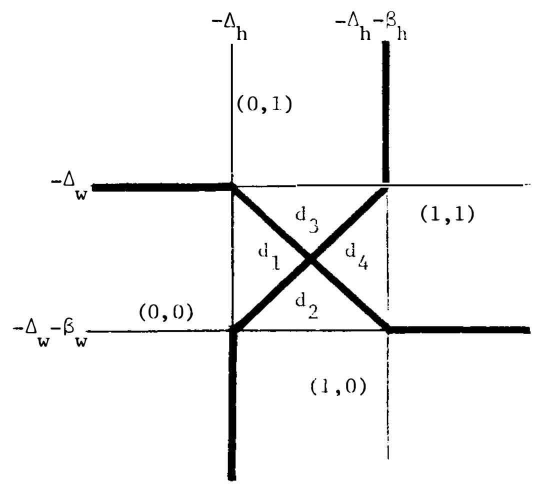

Fig. 3c : $\beta_h < 0$, $\beta_w \geq 0$

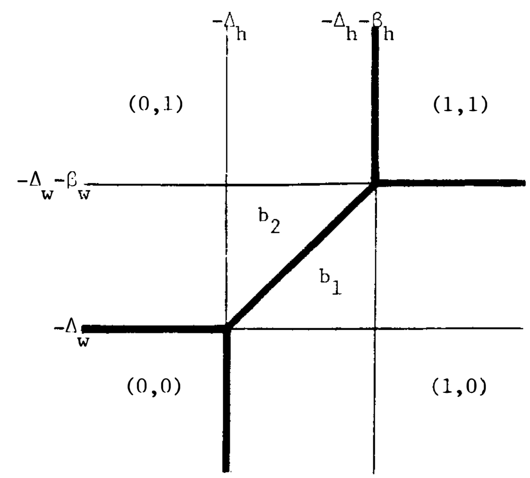

Fig. 3d : $\beta_h < 0$, $\beta_w < 0$

W_2 never occur. Thus:

$$Pr(0,0) = Pr(H_1,W_1),$$
$$Pr(1,0) = Pr(H_3,W_1) + Pr(H_4,W_1) + Pr(H_4,W_3) + b_1Pr(H_3,W_3),$$
$$Pr(0,1) = Pr(H_1,W_3) + Pr(H_1,W_4) + Pr(H_3,W_4) + b_2Pr(H_3,W_3),$$
$$Pr(1,1) = Pr(H_4,W_4).$$

It now suffices to use the conditions on ε_h and ε_w for obtaining particular reaction functions for the husband and wife. For each of the 4 possible pairs of signs for β_h and β_w, figures 3a-3d show the areas over which the bivariate normal density for $(\varepsilon_h,\varepsilon_w)$ must be integrated to obtain the 4 probabilities Pr(0,0), Pr(1,0), Pr(0,1), and Pr(1,1). (The areas are separated by heavy lines, while the lighter lines separate the areas corresponding to the realizations of the specific pairs of reaction functions.)

It follows that the probabillities Pr(0,0), Pr(1,0), Pr(0,1) and Pr(1,1) are given by Equations (8)-(11).

Q.E.D.

c. PROOF OF PROPOSITION 2: Since $(\varepsilon_h,\varepsilon_w)$ have a joint continuous distribution, it follows from the conditions for the husband's and wife's reaction functions (Table 2) that, if $\beta_h = 0$ or $\beta_w = 0$ then H_2, H_3, W_2, or W_3 occur with zero probability.

Conversely, if the pair (H_2,W_2) occurs with zero probability it follows from Table 2 that either $\beta_h \leq 0$ or $\beta_w \leq 0$, i.e., that $\beta_h \cdot \beta_w \geq 0$. Similarly if the pair (H_2,W_3) occurs with zero

probability it follows that $\beta_h \leq 0$ or $\beta_w \geq 0$, i.e., that $\beta_h \cdot \beta_w \leq 0$. Hence $\beta_h \cdot \beta_w = 0$.

Q.E.D.

d. First Partial Derivatives of the Probabilities Pr(i,j)

For brevity, let us rewrite the four probabilities Pr(i,j) listed in PROPOSITION 1 using the indicator variables I_h and I_w defined as:

$$I_h = \begin{cases} 1 & \text{if } \beta_h > 0 \\ 0 & \text{if } \beta_h < 0 \end{cases} \quad \text{and } I_w = \begin{cases} 1 & \text{if } \beta_w > 0 \\ 0 & \text{if } \beta_w < 0 \end{cases}$$

Then we have:

$$Pr(0,0) = F(-\Delta_h, -\Delta_w,\rho) - I_hI_wa_2I_{++} + I_h(1-I_w)c_1I_{+-} + (1-I_h)I_wd_1I_{-+}$$
$$Pr(1,0) = F(\Delta_h, -\Delta_w - \beta_w, -\rho) + I_h(1-I_w)c_2I_{+-} + (1-I_h)I_wd_2I_{-+} - (1-I_h)(1-I_w)b_2I_{--}$$
$$Pr(0,1) = F(-\Delta_h - \beta_h,\Delta_w, -\rho) + I_h(1-I_w)c_3I_{+-} + (1-I_h)I_wd_3I_{-+} - (1-I_h)(1-I_w)b_1I_{--}$$
$$Pr(1,1) = F(\Delta_h + \beta_h,\Delta_w + B_w,\rho) - I_hI_wa_1I_{++} + I_h(1 - I_w)c_4I_{+-} + (1-I_h)I_wd_4I_{-+}$$

Let Φ be the univariate normal c.d.f. and let φ be the corresponding p.d.f. Using the identities $\frac{\partial F(x,y,\rho)}{\partial x} = \varphi(x)\Phi(y^* - \rho x^*)$ and $\frac{\partial F(x,y,\rho)}{\partial y} = \varphi(y)\Phi(x^* - \rho y^*)$, where a quantity with a "*" means that quantity is divided by the square root of $1 - \rho^2$, the first partial derivatives of the probabilities Pr(i,j) are as follows:

$$\frac{\partial F(-\Delta_h, -\Delta_w,\rho)}{\partial\beta_h} = 0 ,$$
$$\frac{\partial F(-\Delta_h, -\Delta_w,\rho)}{\partial\beta_w} = 0 ,$$

$$\frac{\partial F(-\Delta_h, -\Delta_w, \rho)}{\partial \gamma_h} = -\varphi(\Delta_h)\Phi(-\Delta_w^* + \rho\Delta_h^*)x_h ,$$

$$\frac{\partial F(-\Delta_h, -\Delta_w, \rho)}{\partial \gamma_w} = -\varphi(\Delta_w)\Phi(-\Delta_h^* + \rho\Delta_w^*)x_w ;$$

$$\frac{\partial F(\Delta_h, -\Delta_w - \beta_w, -\rho)}{\partial \beta_h} = 0,$$

$$\frac{\partial F(\Delta_h, -\Delta_w - \beta_w, -\rho)}{\partial \beta_w} = -\varphi(\Delta_w + \beta_w)\Phi(\Delta_h^* - \rho(\Delta_w^* + \beta_w^*)),$$

$$\frac{\partial F(\Delta_h, -\Delta_w - \beta_w, -\rho)}{\partial \gamma_h} = \varphi(\Delta_h)\Phi(-\Delta_w^* - \beta_w^* + \rho\Delta_h^*)x_h,$$

$$\frac{\partial F(\Delta_h, -\Delta_w - \beta_w, -\rho)}{\partial \gamma_w} = -\varphi(\Delta_w + \beta_w)\Phi(\Delta_h^* - \rho(\Delta_w^* + \beta_w^*))x_w ;$$

$$\frac{\partial F(-\Delta_h - \beta_h, \Delta_w, -\rho)}{\partial \beta_h} = -\varphi(\Delta_h + \beta_h)\Phi(\Delta_w^* - \rho(\Delta_h^* + \beta_h^*)) ,$$

$$\frac{\partial F(-\Delta_h - \beta_h, \Delta_w, -\rho)}{\partial \beta_w} = 0,$$

$$\frac{\partial F(-\Delta_h - \beta_h, \Delta_w, -\rho)}{\partial \gamma_h} = -\varphi(\Delta_h + \beta_h)\Phi(\Delta_w^* - \rho(\Delta_h^* + \beta_h^*))x_h,$$

$$\frac{\partial F(-\Delta_h - \beta_h, \Delta_w, -\rho)}{\partial \gamma_w} = \varphi(\Delta_w)\Phi(-\Delta_h^* - \beta_h^* + \rho\Delta_w^*)x_w ;$$

$$\frac{\partial F(\Delta_h + \beta_h, \Delta_w + \beta_w, \rho)}{\partial \beta_h} = \varphi(\Delta_h + \beta_h)\Phi(\Delta_w^* + \beta_w^* - \rho(\Delta_h^* + \beta_h^*)) ,$$

$$\frac{\partial F(\Delta_h + \beta_h, \Delta_w + \beta_w, \rho)}{\partial \beta_w} = \varphi(\Delta_w + \beta_w)\Phi(\Delta_h^* + \beta_h^* - \rho(\Delta_w^* + \beta_w^*)) ,$$

$$\frac{\partial F(\Delta_h + \beta_h, \Delta_w + \beta_w, \rho)}{\partial \gamma_h} = \varphi(\Delta_h + \beta_h)\Phi(\Delta_w^* + \beta_w^* - \rho(\Delta_h^* + \beta_h^*))x_h,$$

$$\frac{\partial F(\Delta_h + \beta_h, \Delta_w + \beta_w, \rho)}{\partial \gamma_w} = \varphi(\Delta_w + \beta_w)\Phi(\Delta_h^* + \beta_h^* - \rho(\Delta_w^* + \beta_w^*))x_w ;$$

and

$$\frac{\partial I_{++}}{\partial \beta_h} = -\varphi(\Delta_h + \beta_h)\Phi(-\Delta_w^* - \beta_w^* + \rho(\Delta_h^* + \beta_h^*)) + \varphi(\Delta_h + \beta_h)\Phi(-\Delta_w^* + \rho(\Delta_h^* + \beta_h^*)) ,$$

$$\frac{\partial I_{++}}{\partial \beta_w} = -\varphi(\Delta_w + \beta_w)\Phi(-\Delta_h^* - \beta_h^* + \rho(\Delta_w^* + \beta_w^*)) + \varphi(\Delta_w + \beta_w)\Phi(-\Delta_h^* + \rho(\Delta_w^* + \beta_w^*)) ,$$

$$\frac{\partial I_{++}}{\partial \gamma_h} = [\, -\varphi(\Delta_h)\Phi(-\Delta_w^* + \rho\Delta_h^*) - \varphi(\Delta_h + \beta_h)\Phi(-\Delta_w^* - \beta_w^* + \rho(\Delta_h^* + \beta_h^*))$$
$$+ \varphi(\Delta_h)\Phi(-\Delta_w^* - \beta_w^* + \rho\Delta_h^*) + \varphi(\Delta_h + \beta_h)\Phi(-\Delta_w^* + \rho(\Delta_h^* + \beta_h^*)) \,]\, x_h ,$$

$$\frac{\partial I_{++}}{\partial \gamma_w} = [\, -\varphi(\Delta_w)\Phi(\Delta_h^* + \rho\Delta_w^*) - \varphi(\Delta_w + \beta_w)\Phi(-\Delta_h^* - \beta_h^* + \rho(\Delta_w^* + \beta_w^*))$$
$$+ \varphi(\Delta_w)\Phi(-\Delta_h^* - \beta_h^* + \rho\Delta_w^*) + \varphi(\Delta_w + \beta_w)\Phi(-\Delta_h^* + \rho(\Delta_w^* + \beta_w^*)) \,]\, x_w ,$$

with

$$\frac{\partial I_{+-}}{\partial(.)} = \frac{\partial I_{-+}}{\partial(.)} = -\frac{\partial I_{++}}{\partial(.)} \quad \text{and} \quad \frac{\partial I_{--}}{\partial(.)} = \frac{\partial I_{++}}{\partial(.)} .$$

e. PROOF OF PROPOSITION 3:

Let $Z_t = (Y_{ht}, Y_{wt}, X_{ht}', X_{wt}')$ and $\theta = (\beta_h, \beta_w, \gamma_h, \gamma_w)$. Define:

$$B \equiv E\left[\sum_{t=1}^{T} \frac{\partial \log f(Z_t,\theta)}{\partial \theta} \cdot \frac{\partial \log f(Z_t,\theta)}{\partial \theta'}\right] = \sum_{t=1}^{T} B_t$$

From Section 4, we have, omitting the subscript t, that

$$\frac{\partial \log f(Z,\theta)}{\partial \theta} = \frac{Y_h Y_w}{Pr(1,1)} \frac{\partial Pr(1,1)}{\partial \theta} + \frac{Y_h(1 - Y_w)}{Pr(1,0)} \frac{\partial Pr(1,0)}{\partial \theta}$$
$$+ \frac{(1 - Y_h)Y_w}{Pr(0,1)} \frac{\partial Pr(0,1)}{\partial \theta} + \frac{(1 - Y_h)(1 - Y_w)}{Pr(0,0)} \frac{\partial Pr(0,0)}{\partial \theta}$$

B_t is therefore given by

$$E\begin{bmatrix} \frac{\partial \log f}{\partial \beta_h}\frac{\partial \log f}{\partial \beta_h} & \frac{\partial \log f}{\partial \beta_h}\frac{\partial \log f}{\partial \beta_w} & \frac{\partial \log f}{\partial \beta_h}\frac{\partial \log f}{\partial \gamma_h'} & \frac{\partial \log f}{\partial \beta_h}\frac{\partial \log f}{\partial \gamma_w'} \\ \frac{\partial \log f}{\partial \beta_w}\frac{\partial \log f}{\partial \beta_h} & \frac{\partial \log f}{\partial \beta_w}\frac{\partial \log f}{\partial \beta_w} & \frac{\partial \log f}{\partial \beta_w}\frac{\partial \log f}{\partial \gamma_h'} & \frac{\partial \log f}{\partial \beta_w}\frac{\partial \log f}{\partial \gamma_w'} \\ \frac{\partial \log f}{\partial \gamma_h}\frac{\partial \log f}{\partial \beta_h} & \frac{\partial \log f}{\partial \gamma_h}\frac{\partial \log f}{\partial \beta_w} & \frac{\partial \log f}{\partial \gamma_h}\frac{\partial \log f}{\partial \gamma_h'} & \frac{\partial \log f}{\partial \gamma_h}\frac{\partial \log f}{\partial \gamma_w'} \\ \frac{\partial \log f}{\partial \gamma_w}\frac{\partial \log f}{\partial \beta_h} & \frac{\partial \log f}{\partial \gamma_w}\frac{\partial \log f}{\partial \beta_w} & \frac{\partial \log f}{\partial \gamma_w}\frac{\partial \log f}{\partial \gamma_h'} & \frac{\partial \log f}{\partial \gamma_w}\frac{\partial \log f}{\partial \gamma_w'} \end{bmatrix}$$

For example, $\frac{\partial \log f}{\partial \beta_h}\frac{\partial \log f}{\partial \beta_h}$ is given by

$$\left[\frac{Y_h Y_w}{Pr(1,1)}\frac{\partial Pr(1,1)}{\partial \beta_h}\right]^2 + \left[\frac{Y_h(1-Y_w)}{Pr(1,0)}\frac{\partial Pr(1,0)}{\partial \beta_h}\right]^2$$
$$+ \left[\frac{(1-Y_h)Y_w}{Pr(0,1)}\frac{\partial Pr(0,1)}{\partial \beta_h}\right]^2 + \left[\frac{(1-Y_h)(1-Y_w)}{Pr(0,0)}\frac{\partial Pr(0,0)}{\partial \beta_h}\right]^2$$

where we have used the fact that Y_h and Y_w take on only the values zero or one. Since Y_h and Y_w are random variables where $Y_h = i$, $Y_w = j$ with probability $Pr(i,j)$, $i,j \in \{0,1\}$, we have that

$$E\left[\frac{\partial \log f}{\partial \beta_h}\frac{\partial \log f}{\partial \beta_h}\right] = \frac{1}{Pr(1,1)}\left[\frac{\partial Pr(1,1)}{\partial \beta_h}\right]^2$$
$$+ \frac{1}{Pr(1,0)}\left[\frac{\partial Pr(1,0)}{\partial \beta_h}\right]^2 + \frac{1}{Pr(0,1)}\left[\frac{\partial Pr(0,1)}{\partial \beta_h}\right]^2 + \frac{1}{Pr(0,0)}\left[\frac{\partial Pr(0,0)}{\partial \beta_h}\right]^2$$

Proceeding analogously, the remaining terms in B are given by:

$$E\left[\frac{\partial \log f}{\partial \theta_k}\cdot\frac{\partial \log f}{\partial \theta_h}\right] = \sum_{i=0}^{1}\sum_{j=0}^{1}\frac{1}{Pr(i,j)}\frac{\partial Pr(i,j)}{\partial \theta_k}\frac{\partial Pr(i,j)}{\partial \theta_h}.$$

Notice that B can be decomposed into $B = A'DA$ where A is of dimension 4T by K, $K = k_h + K_w + 2$, that has as its t-th block:

$$\begin{bmatrix} \frac{\partial Pr_t(1,1)}{\partial \beta_h} & \frac{\partial Pr_t(1,1)}{\partial \beta_w} & \frac{\partial Pr_t(1,1)}{\partial \gamma_h'} & \frac{\partial Pr_t(1,1)}{\partial \gamma_w'} \\ \frac{\partial Pr_t(1,0)}{\partial \beta_h} & \frac{\partial Pr_t(1,0)}{\partial \beta_w} & \frac{\partial Pr_t(1,0)}{\partial \gamma_h'} & \frac{\partial Pr_t(1,0)}{\partial \gamma_w'} \\ \frac{\partial Pr_t(0,1)}{\partial \beta_h} & \frac{\partial Pr_t(0,1)}{\partial \beta_w} & \frac{\partial Pr_t(0,1)}{\partial \gamma_h'} & \frac{\partial Pr_t(0,1)}{\partial \gamma_w'} \\ \frac{\partial Pr_t(0,0)}{\partial \beta_h} & \frac{\partial Pr_t(0,0)}{\partial \beta_w} & \frac{\partial Pr_t(0,0)}{\partial \gamma_h'} & \frac{\partial Pr_t(0,0)}{\partial \gamma_w'} \end{bmatrix}$$

and D is a block diagonal matrix of order 4T, the t-th block given by

$$\begin{bmatrix} Pr_t(1,1) & 0 & 0 & 0 \\ 0 & Pr_t(1,0) & 0 & 0 \\ 0 & 0 & Pr_t(0,1) & 0 \\ 0 & 0 & 0 & Pr_t(0,0) \end{bmatrix}^{-1}$$

The model will be identified if and only if B is nonsingular (see, e.g., Rothenberg (1971) and Bowden (1973)). Since D is of full rank and $4T > K$, a necessary and sufficient condition is that A have full column rank. From part (d) of Appendix A, it is seen that the partial derivatives of $Pr_t(i,j)$ with respect to the vector θ depend on the signs of β_h and β_w; we must therefore check that matrix A is nonsingular for all cases.

Case 1: $\beta_h > 0$, $\beta_w > 0$

Substituting into A_t the partial derivatives, using the notation $e_t^i, f_t^i, g_t^i, h_t^i, i = h,w$, established in the text, matrix A_t can be written as

$$\begin{bmatrix} e^h_t - a_1(e^h_t+h^h_t) & e^w_t - a_1(e^w_t+g^w_t) & [e^h_t - a_1(e^h_t - f^h_t + g^h_t + h^h_t)]x'_{ht} & [e^w_t - a_1(e^w_t - f^w_t + g^w_t + h^w_t)] \\ 0 & g^w_t & g^h_t x'_{ht} & g^w_t x'_{wt} \\ h^h_t & 0 & h^h_t x'_{ht} & h^w_t x'_{wt} \\ -a_2(e^h_t + h^h_t) & -a_2(e^w_t+g^w_t) & [-f^h_t - a_2(e^h_t - f^h_t + g^h_t + h^h_t)]x'_{ht} & [-f^w_t - a_2(e^w_t - f^w_t + g^w_t + h^w_t)]x'_{wt} \end{bmatrix}$$

Now perform the following matrix algebra

(i) subtract Row 4 from Row 1, noting that $a_1 = a_2 = \frac{1}{2}$

(ii) add 1/2 of Row 1 to Row 4

(iii) adding 1/2(Row 2 + Row 3) to Row 4.

(iv) switch Columns 2 and 3

Omitting Row 4 since it is identically null, we have

$$\bar{A}_t = \begin{bmatrix} e^h_t & (e^h_t + f^h_t)x'_{ht} & e^w_t & (e^w_t + f^w_t)x'_{wt} \\ 0 & g^h_t x'_{ht} & g^w_t & g^w_t x'_{wt} \\ h^h_t & h^h_t x'_{ht} & 0 & h^w_t x'_{wt} \end{bmatrix}$$

We now decompose the resulting matrix $\tilde{A}$ into a partitioned matrix

$$\bar{A} = \left[D_h \bar{X}_h \;\middle|\; D_w \bar{X}_w \right]$$

where D_h and D_w are each block diagonal matrices of order 3T, the t-th block of D_h and D_w being D_{ht} and D_{wt} respectively, as given in the text. Since D_h and D_w are both nonsingular, for purposes of identification $\bar{A}$ may be rewritten as

$$\tilde{A} = \left[\bar{X}_h \;\middle|\; D_h^{-1} D_w \bar{X}_w \right]$$

Case 2: $B_h < 0$, $B_w < 0$

Substituting into A_t the partial derivatives found in Appendix A, again using the e^i_t, f^i_t, g^i_t, h^i_t, i = h,w, as established in the text, now perform the following matrix algebra on matrix A

(i) subtract 1/2 of Row 2 from Row 3 noting that $b_1 = b_2 = 1/2$

(ii) add 1/2(Row 1 + Row 3 + Row 4) to Row 2

(iii) multiply Row 2 by two, adding it to Row 4

(iv) reverse Columns 2 and 3

Again, omitting Row 4 since it is identically null, we have

$$\bar{A}_t = \begin{bmatrix} e^h_t & e^h_t x'_{ht} & e^w_t & e^w_t x'_{wt} \\ 0 & f^h_t x'_{ht} & 0 & f^w_t x'_{wt} \\ h^h_t & (h^h_t - g^h_t)x'_{ht} & -g^w_t & (h^w_t - g^w_t)x'_{w} \end{bmatrix}$$

which can be written as $\tilde{A}$.

Case 3: $\beta_h > 0$, $\beta_w < 0$

Proceeding as in Case 2, perform the following algebra

(i) subtract Row 4 from Rows 1, 2 and 3, noting that $d_1 = d_2 = d_3 = d_4 = \frac{1}{4}$

(ii) add 1/4(Row 1 + Row 2 + Row 3) to Row 4

(iii) reverse Columns 2 and 3

Deleting Row 4 since it is null, we have

$$\bar{A}_t = \begin{bmatrix} e_t^h & (e_t^h + f_t^h)x_{ht}' & e_t^w & (e_t^w + f_t^w)x_{wt}' \\ 0 & (g_t^h + f_t^h)x_{ht}' & g_t^w & (g_t^w + f_t^w)x_{wt}' \\ h_t^h & (h_t^h + f_t^h)x_{ht}' & 0 & (h_t^w + f_t^w)x_{wt}' \end{bmatrix}$$

which can be written as $\tilde{A}$.

Case 4: $\beta_h < 0,\ \beta_w > 0$

Identical to Case 3, this can be easily seen by noting in Part (d) of Appendix A that $c_i = d_i = 1/4,\ i = 1,2,3,4$.

Q.E.D.

PROOF OF COROLLARY 1:

Case 1: $\beta_h > 0,\ \beta_w > 0$

Since $x_{ht}' = x_{wt}' \equiv x_t'$ and $\gamma_h = \gamma_w$, it follows that $\Delta_{ht} = \Delta_{wt} \equiv \Delta_t$. Since $\beta_h = \beta_w \equiv \beta$, it further follows that $e_t^h = e_t^w \equiv e_t > 0$, $f_t^h = f_t^w \equiv f_t > 0$, $g_t^h = h_t^w \equiv g_t > 0$, and $h_t^h = g_t^w \equiv h_t < 0$. From the proof of Proposition 3, it follows that $\bar{A}_t$ can be written as

$$\bar{A}_t = \begin{bmatrix} e_t & (e_t + f_t)x_t' & e_t & (e_t + f_t)x_t' \\ 0 & g_t x_t' & h_t & h_t x_t' \\ h_t & h_t x_t' & 0 & g_t x_t' \end{bmatrix}$$

Now perform the following sequence of steps to matrix $\bar{A}_t$. Subtract Column 4 from Column 2; subtract Column 3 from Column 1; add Row 3 to Row 2; multiply Row 3 by -2; add Row 2 to Row 3; add 1/2 of Column 2 to Column 4; multiply Row 3 by $-1/2$; add 1/2 Column 1 to Column 3; divide Row 1 by $(e_t + f_t)$, $(e_t + f_t \neq 0)$; finally, divide Row 3 by $(h_t - g_t)$, $(h_t - g_t \neq 0)$. Calling this new matrix $\tilde{A}_t$, we have

$$\tilde{A}_t = \begin{bmatrix} 0 & 0 & e_t/(e_t + f_t) & x_t' \\ 0 & 0 & h_t & (h_t + g_t)x_t' \\ h_t/(h_t - g_t) & x_t' & 0 & 0 \end{bmatrix} \quad \text{where}$$

rank $\bar{A}$ = rank $\tilde{A}$ = rank M_1 + rank M_2, M_1 and M_2 being T by (K+1) and 2T by (K+1), respectively. The t-th element of M_1 is given by $h_t/(h_t - g_t) \cdot x_t'$; the t-th element of M_2 given

$$\text{by} \quad \begin{bmatrix} e_t/(e_t + f_t) & x_t' \\ h_t & (h_t + g_t)x_t' \end{bmatrix}$$

We will now show by contradiction that M_2 is of full column rank. Assume not. Then there exists $\lambda_0, \lambda_1, \ldots, \lambda_K \neq (0,0,\ldots,0)$ such that for all t

$$\lambda_0 \frac{e_t}{(e_t + f_t)} + \lambda_1 x_{1t} + \cdots \lambda_K x_{Kt} = 0$$

and

$$\lambda_0 h_t + (h_t + g_t)[\lambda_1 x_{1t} + \cdots + \lambda_K x_{Kt}] = 0$$

But, by assumption, Matrix X is of full column rank so that $\lambda_0 \neq 0$. Using the above two equations, we get

$h_t - e_t(h_t + g_t)/(e_t + f_t) = 0$ or $e_t g_t = h_t f_t$ for all t.

But $e_t g_t > 0$ and $h_t f_t < 0$. Contradiction. As is quite clear, the occurrence of $h_t/(h_t - g_t)$ as a linear of x_t', for all t, is extremely small.

Case 2: $\beta_h < 0$, $\beta_w < 0$

Case 2 proceeds similarly to Case 1. Again, using the proof of Proposition 3, $\bar{A}_t$ can be written as

$$\bar{A}_t = \begin{bmatrix} e_t & e_t x_t' & e_t & e_t x_t' \\ 0 & f_t x_t' & 0 & f_t x_t' \\ h_t & (h_t - g_t)x_t' & -h_t & (g_t - h_t)x_t' \end{bmatrix}$$

Now perform the following sequence of operations on $\bar{A}_t$. Subtract Column 1 from Column 3; subtract Column 2 from Column 4; add 1/2 of Column 4 to Column 2; add 1/2 of Column 3 to Column 1; divide Row 3 by -2; divide Row 1 by e_t; divide Row 2 by f_t; finally, divide Row 3 by $(h_t - g_t)$, $(h_t - g_t \neq 0)$. Again, calling this new matrix $\tilde{A}_t$ we have

$$\tilde{A}_t = \begin{bmatrix} 1 & x_t' & 0 & 0 \\ 0 & x_t' & 0 & 0 \\ 0 & 0 & h_t/(h_t - g_t) & x_t' \end{bmatrix} \text{ where}$$

rank $\bar{A}$ = rank $\tilde{A}$ = rank M_1 + rank M_2, M_1 and M_2 dimensioned the same as Case 1. The t-th element of M_1 is given by $h_t/(h_t - g_t)$, as in Case 1; the t-th element of M_2 is given by

$\begin{bmatrix} 1 & x_t' \\ 0 & x_t' \end{bmatrix}$. By inspection, M_2 has full column rank.

Q.E.D.

PROOF OF COROLLARY 2:

Note that when only constant terms are included, $\bar{X}_h$ and $\bar{X}_w$ reduce respectively to T repetitions of

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Note also that the blocks comprising D_h are now identical, as are the blocks comprising D_w now identical. (This is because $\Delta_{ht}, \Delta^*_{ht}, \Delta_{wt}$, and Δ^*_{wt} no longer depend on t). As a result, the matrix $\tilde{A}$ repeats itself every three rows. Since K = 4 > 3, B is singular.

Q.E.D.

LEMMA 1: For a two-person, two strategy, normal form game, if there is a unique Nash Equilibrium in pure strategies, then there are no Nash Equilibria in mixed strategies.

PROOF: (by contradiction) Let the payoff matrix be given by:

		player B	
		1′	2′
player A	1	a_{11}, b_{11}	a_{12}, b_{12}
	2	a_{21}, b_{21}	a_{22}, b_{22}

Without loss of generality, assume (1,1) is the unique Nash Equilibrium; we therefore have $a_{11} \geq a_{21}$ and $b_{11} \geq b_{12}$. Assume a mixed strategy N.E. exists where player A plays strategy 1 with probability p and strategy 2 with probability (1-p), $0 < p < 1$. Similarly, player B plays strategy 1′ with probability q and strategy 2′ with probability (1-q), $0 < q < 1$. In choosing mixed strategies, players A and B must solve respectively:

$$\max_p E\pi_A = pqa_{11} + p(1 - q)a_{12} + (1 - p)qa_{21} + (1 - p)(1 - q)a_{22}$$

$$\max_q E\pi_B = pqb_{11} + p(1 - q)b_{12} + (1 - p)qb_{21} + (1 - p)(1 - q)b_{22}$$

For a mixed strategy N.E. to exist, it must be the case that:

$$\frac{\partial E\pi_A}{\partial p} = 0 \Rightarrow a_{11}q + a_{12}(1 - q) = a_{21}q + a_{22}(1 - q) \qquad (1)$$

and

$$\frac{\partial E\pi_B}{\partial q} = 0 \Rightarrow b_{11}p + b_{21}(1 - p) = b_{12}p + b_{22}(1 - p) \qquad (2)$$

Case 1: $a_{11} = a_{21} \Rightarrow a_{12} = a_{22}$ by (1)

$b_{11} = b_{12} \Rightarrow b_{21} = b_{22}$ by (2)

We have that both (1,1) and (2,2) are N.E. Contradiction

Case 2: $a_{11} > a_{21} \Rightarrow a_{12} < a_{22}$ by (1)

$b_{11} = b_{12} \Rightarrow b_{21} = b_{22}$ by (2)

Both (1,1) and (2,2) are N.E. Contradiction

Case 3: $a_{11} = a_{21} \Rightarrow a_{12} = a_{22}$ by (1)

$b_{11} > b_{12} \Rightarrow b_{21} < b_{22}$ by (2)

Both (1,1) and (2,2) are N.E. Contradiction

Case 4: $a_{11} > a_{21} \Rightarrow a_{12} < a_{22}$ by (1)

$b_{11} > b_{12} \Rightarrow b_{21} < b_{22}$ by (2) Contradiction

Q.E.D.

LEMMA 2: For a two-person, two strategy, normal form game, if there exists two pure strategy Nash Equilibria, in addition to a mixed strategy N.E., then the mixed strategy is dominated by either

(1) one of the two pure strategy Nash Equilibria or

(2) a cooperative game solution.

PROOF: Let the payoff matrix for Players A and B be the same as that given in the proof of Lemma 1. Assume two pure strategy Nash Equilibria; without loss of generality, let them be (1,1′) with payoff (a_{11}, b_{11}) and (2,2′) with payoff (a_{22}, b_{22}). For a mixed strategy N.E. to exist, we know from Lemma 1 that we must have

$$\frac{\partial E\pi_A}{\partial p} = a_{11}q - a_{21}q + a_{12}(1 - q) - a_{22}(1 - q) = 0 \qquad (1)$$

$$\Rightarrow q = \frac{a_{22} - a_{12}}{a_{11} - a_{21} - a_{12} + a_{22}} \quad \text{with} \quad 0 < q < 1 \qquad (2)$$

and

$$\frac{\partial E\pi_B}{\partial q} = b_{11}p + b_{21}(1 - p) - b_{12}p - b_{22}(1 - p) = 0 \tag{3}$$

$$\Rightarrow p = \frac{b_{22} - b_{21}}{b_{11} - b_{12} + b_{22} - b_{21}} \quad \text{with} \quad 0 < p < 1 \tag{4}$$

Necessary and sufficient conditions for (2) and (4) to hold are that

$$\begin{aligned} &a_{22} - a_{12} > 0, \; b_{22} - b_{21} > 0 \\ &a_{11} - a_{21} > 0, \; b_{11} - b_{12} > 0 \end{aligned} \tag{5}$$

Substituting (2) into (1), we get the payoff for player A:

$$\pi_A = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}}$$

Similarly, substituting (4) into (2) gives the payoff for player B:

$$\pi_B = \frac{b_{11}b_{22} - b_{12}b_{21}}{b_{11} - b_{12} + b_{22} - b_{21}}.$$

Therefore the payoff to players A and B when they play a mixed strategy is given by (π_A, π_B). For ease of notation, let $(C_1, C_2) \equiv (\pi_A, \pi_B)$. We want to compare (a_{11}, b_{11}) and (a_{22}, b_{22}) with (C_1, C_2).

Now $$a_{11} - C_1 = \frac{(a_{11} - a_{12})(a_{11} - a_{21})}{a_{11} + a_{22} - a_{12} - a_{21}}$$

and

$$a_{22} - C_2 = \frac{(a_{22} - a_{21})(a_{22} - a_{12})}{a_{11} + a_{22} - a_{12} - a_{21}}.$$

From (5), we have that

$$\text{sgn}(a_{11} - C_1) = \text{sgn}(a_{11} - a_{12})$$

$$\text{sgn}(a_{22} - C_1) = \text{sgn}(a_{22} - a_{21})$$

If $a_{12} > a_{11}$, then $a_{11} < C_1$. Also, $a_{22} > a_{12} > a_{11} > a_{21} \Rightarrow a_{22} > C_1$.

If $a_{22} > a_{21}$, then $a_{22} < C_1$. Also, $a_{11} > a_{21} > a_{22} > a_{12} \Rightarrow a_{11} > C_1$.

In either case, $\max\{a_{11}, a_{22}\} > C_1$.

Also, $b_{11} - C_2 = \dfrac{(b_{11} - b_{21})(b_{11} - b_{12})}{b_{11} + b_{22} - b_{12} - b_{21}}$ and

$b_{22} - C_2 = \dfrac{(b_{22} - b_{21})(b_{22} - b_{12})}{b_{11} + b_{22} - b_{21} - b_{12}}$.

Again from (5), we have that

$$\text{sgn}(b_{11} - C_2) = \text{sgn}(b_{11} - b_{21})$$

$$\text{sgn}(b_{22} - C_2) = \text{sgn}(b_{22} - b_{12})$$

If $b_{11} < b_{21}$, then $b_{11} < C_2$. Also, $b_{22} > b_{21} > b_{11} > b_{12} \Rightarrow b_{22} > C_2$.

If $b_{22} < b_{12}$, then $b_{22} < C_2$. Also, $b_{11} > b_{12} > b_{22} > b_{21} \Rightarrow b_{11} > C_2$.

In either case, $\max\{b_{11}, b_{22}\} > C_2$. Now if $\max\{a_{11}, a_{22}\} = a_{11}$ and $\max\{b_{11}, b_{22}\} = b_{11}$ or $\max\{a_{11}, a_{22}\} = a_{22}$ and $\max\{b_{11}, b_{22}\} = b_{22}$, then we know that the mixed strategy N.E. is dominated by a pure strategy. On the other hand, if $\max\{a_{11}, a_{22}\} = a_{11}$ and $\max\{b_{11}, b_{22}\}$ b_{22} or $\max\{a_{11}, a_{22}\} = a_{22}$ and $\max\{b_{11}, b_{22}\} = b_{11}$, we revert to a bargaining solution from cooperative game theory, which dominates a mixed strategy Nash Equilibrium. In either case, the mixed strategy will be dominated by some game theoretic solution.

Q.E.D.

LEMMA 3: For a two-person, two strategy, normal form game, if there are no Nash Equilibria in pure strategies, then there exists a unique Nash Equilibrium in mixed strategies. Existence is well known (see e.g., Owen (1982), p. 126). From Lemma 1 we know that for a mixed strategy N.E. to exist, we must have

$$\frac{\partial E\pi_A}{\partial p} = 0 \Rightarrow q = \frac{a_{22}-a_{12}}{a_{11}-a_{12}-a_{21}+a_{22}},\quad a_{11}-a_{12}-a_{21}+a_{22} \neq 0 \qquad (1)$$

and

$$\frac{\partial E\pi_B}{\partial q} = 0 \Rightarrow p = \frac{b_{22}-b_{21}}{b_{11}-b_{12}-b_{21}+b_{22}},\quad b_{11}-b_{12}-b_{21}+b_{22} \neq 0 \qquad (2)$$

We need only show that $q \neq 0$, $q \neq 1$, $p \neq 0$, and $p \neq 1$. Assume $q = 0 \Rightarrow a_{22} = a_{12} \Rightarrow$ a pure strategy N.E. exists. Assume $q = 1 \Rightarrow a_{11} = a_{21} \Rightarrow$ a pure strategy N.E. exists. Similarly, for $p = 0$ and $p = 1$, a pure strategy N.E. exists. Contradiction.

Q.E.D.

APPENDIX B

MEANS AND VARIANCES OF THE EXPLANATORY VARIABLES

		MEAN	VARIANCE
AGEH	Age of Husband	36.7	109.0
AGEW	Age of Wife	34.3	97.4
EDUCH	Husband's education (years)	12.6	4.32
EDUCW	Wife's education (years)	12.7	6.65
UNEM	Local uemployment rate	9.49	13.8
ASSETS	Family non-wage income (in thousands $)	.986	22.6
KIDS13	Number of children 13 or younger	1.22	1.28
KIDS14	Number of children 14 or older	.296	.458
RACE	Race dummy (1 if Black or Hispanic, 0 otherwise	.252	-

Below we list the selection criteria used. Of the original 6742 observations, we were left with 2020 observations after selection.

(1) Family composition change in 1982 was restricted to children moving in or out of the home; husband and wife remained married and in the home. Loss: 2295 observations.

(2) Family was restricted to husband, wife and children. Loss: 1349 observations.

(3) Both the husband and the wife were restricted to be 64 years or less in age. Loss: 219 observations.

(4) Husbands who stated they were retired, permanently disabled, temporarily laid-off, or students were excluded. Those who stated they were "working now" were classified by us as working; those who answered "looking for work, unemployed," were considered to be not working. Loss: 331 observations.

(5) Wives who stated they were retired, permanently disabled, temporarily laid-off, unemployed but looking for work, or students were excluded from the sample. Those who stated they were "working now" were classified by us as working; those who answered "housewife" were considered to be not working. Loss: 174 observations.

(6) If either the husband or the wife had a physical or nervous condition that limited the type or the amount of work they could do, they were excluded from the sample. Loss: 287 observations.

(7) If any record contained missing data for the 10 explanatory variables used in the analysis, that record was dropped. Loss: 67 observations.

Of the 2020 observations remaining after selection, the numerical breakdown based upon husband/wife employment status is described in the following table.

		HUSBAND	
		Working	Not working
WIFE	Working	1204	48
	Not working	739	29

FOOTNOTES

*. The early stage of this work was done while the second author was visiting the Universite' des Sciences Sociales de Toulouse where he benefited from insightful discussions with J. J. Laffont. D. Lien and R. McKelvey provided helpful comments on the game theoretic part of this paper. Finally, we are greatly indebted to D. Grether and D. Rivers for valuable criticisms. Remaining errors are of course ours.

1. When an individual is indifferent between working and not working, we arbitrarily considers that he will work, hence the use of the weak inequality.

2. The condition $\beta_h = 0$ holds if one assumes that the husband's utility function is additively separable so that $U_h(Y_h, Y_w) = U_h(Y_h) + U_w(Y_w)$. Then $\Delta_h = U_h(1) - U_h(0)$. I am grateful to Donald Lien for pointing this out. Note, however, that assuming the husband's utility function to be additively separable essentially removes simultaneity as is argued in the text.

3. See, however, Vuong (1980, 1981, 1982).

4. An intuitive argument for such a result is that, when there are no explanatory variables other than the two constant terms, then the model has four parameters (including β_h and β_w) while there are four probabilities to be explained of which only three are

independent.

5. The set of explanatory variables used in our empirical analysis will present no surprises; Indeed, most empiriccal studies of labor force participation using cross-section data use a fairly common set of explanatory variables (see,e.g., Ashenfelter and Heckman (1974), Gronau (1973), Heckman (1974,1976), and Nakamura and Nakamura (1981)).

6. The Panel Study of Income Dynamics asked only the race of the head of household; if married, the Panel Study assumed the spouse to be of the same race.

7. For a discussion of the appropriateness of including current assets in a labor supply equation, see Cotterman (1981).

8. Note from (30) that $\tilde{\gamma}_w^5$ and $\tilde{\gamma}_w^6$ enter with negative signs.

9. See, e.g. Lazear (1977) for a discussion of this particular hypothesis.

REFERENCES

Amemiya, T. "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables Are Truncated Normal." Econometrica 42 (1974):999-1012.

Ashenfelter, O. and Heckman, J. "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply." Econometrica 42 (1974):73-85.

Berndt, E. R.; Hall, B. H.; Hall, R. E. and Hausman, J. A. "Estimation and Inference in Non-Linear Structural Models." Annals of Economic and Social Measurement 3 (1974):653-665.

Bowden, R. "The Theory of Parametric Identification." Econometrica 41 (1973):1069-1074.

Cotterman, R. F. "The role of Assets in Labor Supply Functions." Economic Inquiry 19 (1981):495-505.

Goodman, L. A. "The Multivariate Analysis of Qualitative Data: Interactions Among Multiple Classifications." Journal of the American Statistical Association 65 (1970):226-256.

Gronau, R. "The Effect of Children on the Housewife's Value of Time." Journal of Political Economy Supplement, 81 (1973):168-199.

Gourieroux, C.; Laffont, J. J. and Monfort, A. "Coherency Conditions in Simultaneous Linear Equations Models with Endogenous Switching Regimes." Econometrica 48 (1980):675-695.

Haberman, S. J. The Analysis of Frequency Data. Chicago: University of Chicago Press, 1974.

Heckman, J. "Shadow Prices, Market Wages, and Labor Supply." Econometrica 42(1974):679-694.

__________. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for such Models." Annals of Economic and Social Measurement 5(1976):475-492.

__________. "Dummy Endogenous Variables in a Simultaneous Equation System." Econometrica 46 (1978):931-959.

Heckman, J. and Macurdy, T. "A Life Cycle Model of Female Labour Supply." Review of Economic Studies 47(1980):47-74.

Koopmans, T. C. and Hood, W. C. Studies in Econometric Method. New York: Wiley, 1953.

Lazear, E. "Education: Consumption or Production?" Journal of Political Economy 85(1977):569-597.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior." In Frontiers of Econometrics, edited by P. Zarembka. New York: Academic Press, 1974.

_________. "Econometric Models of Probabilistic Choice." In Structural Analysis of Discrete Data with Econometric Applications, edited by C. Manski and D. McFadden. Cambridge: Massachusetts Institute of Technology Press, 1981.

Maddala, G. S. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge University Press, 1983.

Maddala, G. S. and Lee, L. F. "Recursive Models with Qualitative Endogenous Variables." Annals of Economic and Social Measurement 5 (1976):525-545.

Nakamura, A. and Nakamura, M. "A Comparison of the Labor Force Behavior of Married Women in the United States and Canada, with Special Attention to the Impact of Income Taxes." Econometrica 49(19811):451-489.

Nelson, F. D. and Olsen, L. "Specification and Estimation of a Simultaneous Equation Model with Limited Dependent Variables." International Economic Review 19 (1978):695-710.

Nerlove, M. and Press, S. J. Univariate and Multivariate Log-Linear and Logistic Models. Santa Monica: Rand

Corporation, 1973.

_________. "Multivariate Log-Linear Probability Models for the Analysis of Qualitative Data." Discussion Paper no. 1. Center for Statistics and Probability, Northwestern University, 1976.

Owen, G. Game Theory. New York: Academic Press, 1982.

Rothenberg, T. J. "Identification in Parametric Models." Econometrica 39 (1971):577-591.

Schmidt, P. "Constraints on the Parameters in Simultaneous Tobit and Probit Models." In Structural Analysis of Discrete Data with Econometric Applications, edited by C. Manski and D. McFadden. Cambridge: Massachusetts Institute of Technology Press, 1981.

Vuong, Q. H. "Heckman's Criticism of the Nerlove-Press Model: A Reply." Mimeographed paper, 1980.

_________. "Structural log-linear Probability Models." Working Paper 8114, GREMAQ, Universite' des Sciences Sociales, Toulouse 1981.

_________. "Conditional Log-Linear Probability Models: A Theoretical Development with an Empirical Application." Ph.D. dissertation, Northwestern University, 1982.

________. "Probability Feedback in a Recursive System of Probability Models." Social Science Working Paper no. 443. Pasadena: California Institute of Technology, September 1982.

Waldman, D. M. "An Economic Interpretation of Parameter Constraints in a Simultaneous Equations Model with Limited Dependent Variables." International Economic Review 22 (1981):731-739.